Coming to Your Senses: Promoting Critical Thinking about Sensors through Playful Interaction in Classrooms

Susan Lechelt
University of Edinburgh
Edinburgh, United Kingdom
zlechelt@ed.ac.uk

Yvonne Rogers
University College London
London, United Kingdom
y.rogers@ucl.ac.uk

Nicolai Marquardt
University College London
London, United Kingdom
n.marquardt@ucl.ac.uk

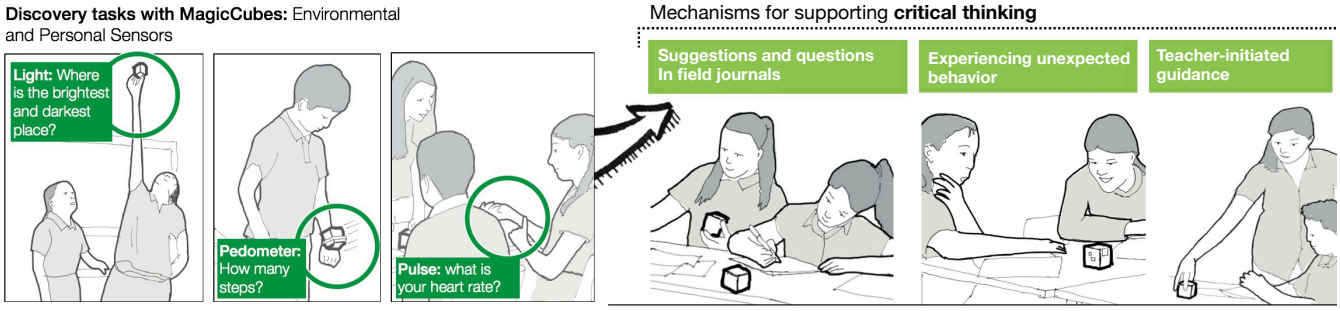

Figure 1: Overview of the playful, exploration-based tasks, and the mechanisms for supporting critical thinking.

ABSTRACT

Learning through exploration is assumed to be a powerful way of introducing children to computer science concepts. However, it is uncertain how exploring physical computing toolkits can promote movement between conceptual knowledge and abstract reflection, and lead to critical thinking about technology. We investigated how children aged 9-11 years explored and reasoned about personal and environmental data sensors, using a playful exploration-based physical toolkit in their classroom. We report on the ways in which critical thinking about sensor accuracy and reliability developed through reflective dialogue and playful interaction, taking into account the support structures embedded in the classroom. Finally, we discuss strategies for designing exploration-based learning for classroom settings, to promote critical thinking about data sensing.

Author Keywords

Computer science education, critical thinking, embodied interaction, collaborative learning

CSS Concepts

• **Human-centered computing~Human computer interaction (HCI)** • **Social and professional topics~Computing education** • Social and professional topics~Children

Submitted to ACM *IDC '20,* June 21–24, 2020, London, United Kingdom

INTRODUCTION

While there is much research about how to teach computational thinking in schools through making and programming [23,27,28,34], equally important is how to introduce *critical thinking* about computing. Critical thinking is becoming an increasingly core component of primary and secondary computing curricula [36,38] and has been called one of the key skills for the 21st century by leaders in business, education and policy [24]. However, learning to think critically about technology is not straightforward. It requires students to go beyond just understanding computational concepts, and instead learn to reason about and evaluate the benefits and limitations of specific technologies. We are interested in the question of how to promote this type of thinking in young children, in particular those at the end of primary school. The reason we focus on this age group is that it is when children start experimenting with and learning about technology as part of the UK national computing curriculum [36].

The focus of our research is to investigate how to enable children to critically reflect on the processes of *sensing and collecting data*, which are core aspects of the Internet of Things (IoT) and ubiquitous computing. Different sensors collect data differently, and some are more accurate, reliable and informative than others. For example, a pedometer's accuracy is contingent on how a step is defined, as well as where the pedometer is placed. Equally, how informative a galvanic skin response value is, depends on the context of use. Given the increasing ubiquity of sensors - and IoT more broadly - in our everyday interactions with technology, we view critical thinking about data collection and sensing to be an important skill for all children to learn. However, little research has so far been carried out on appropriate methods

to teach school-aged children to reflect on how different kinds of sensor data are collected, what their reliability and accuracy are, and to what extent the data can be trusted. Our approach is to design learning activities in conjunction with using a physical computing toolkit that can enable children to move between *learning new concepts* about sensors and *thinking critically* about their limitations and context of use.

It is one thing to be taught the definitions of high-level technology concepts like reliability and accuracy; it is another to be able to put them into practice and operationalize them for different problem spaces. Therefore, teaching this level of understanding and application requires considering how learning tasks can be developed to foster curiosity, experimentation, and importantly, "stepping out" [2] from the situated activity in order to analyze, evaluate, and reflect on what is being learned. As part of our approach, we used a sensor-based physical toolkit, called the Magic Cubes [39], that has been found to enable children to readily explore and experiment with sensing, measuring, and collecting data [18]. Here we wanted to see whether children could also use the toolkit to learn how to evaluate the reliability and accuracy of data and how informative it is. The context we chose for this was data they could collect using the toolkit to sense aspects of their bodies and their environment. We conducted a series of classroom studies for specific learning activities. The studies were aimed at addressing the following research question:

Can exploration-based physical toolkits enable children to think critically about sensors, sensing and sensor data in a classroom context? If so, to what extent and how?

Our findings were revealing, showing how the mechanics of reflection on sensors, sensing and sensor properties can be triggered and applied during learning sessions in classrooms. We analyze and discuss how the sensor properties, the pedagogical materials and the instructors were instrumental in facilitating critical thinking.

BACKGROUND

What is critical thinking?

Since the 1980s, a wide body of literature has been concerned with defining what specific cognitive processes and skills comprise critical thinking (e.g., [8,11,22]). Although critical thinking has been defined and operationalized in a variety of ways, key researchers in the domain, including Ennis [8], Facione [9] and Halpern [12] have agreed that it comprises a number of key abilities. Lai [17] summarises these as including: analyzing arguments, claims or evidence; making inferences using inductive or deductive reasoning; and evaluating and making decisions or solving problems. There is still much debate about the skills and processes encapsulated within critical thinking, and to what extent they are observable and generalizable beyond a specific context (e.g., [4,5]). Here, based on previous definitions of critical thinking, our research is constrained to putative cognitive processes that can be viewed as *useful* to learning about sensors and sensing. Specifically, by mapping the key tenets of critical thinking as summarized by Lai [17] to the domain sensing, we decided to focus this study on investigating how children can be supported in the processes of:

(i) **understanding** what sensors measure and how
(ii) **observing**, **experimenting with** and **analyzing** representations of sensor data to reason about when a sensor may not be working as expected and why
(iii) and **evaluating** information gathered about sensors in order to reason about how reliable, accurate and informative they are in general.

Designing for critical thinking

Despite critical thinking about computing and technology being considered increasingly important, there has been little work on teaching it as an explicit outcome in classrooms. The focus has largely been on problem solving when making or programming (e.g., [7,35]), rather than reflecting on using technology critically. This suggests that there is still a need to consider how to teach critical thinking skills, which enable children to be able to judge the merits and pitfalls of a type of technology in relation to how it fits into everyday life.

How then to design for the cognitive processes and behaviors that underlie critical thinking? Because critical thinking transcends declarative knowledge [17] it is not enough to teach it just by lecturing. A number of researchers have argued that it is best taught by combining constructivist learning-by-doing [25] and social constructivist collaborative learning [1,37]. Specifically, learning through a constructivist, grounded experience may promote critical thinking processes like applying knowledge to new contexts, analyzing the problem at hand and making explicit or implicit inferences [17]. Moreover, it has been suggested that designing authentic learning tasks that are meaningful to the learner can promote reflection [3,26].

While active, constructive exploration may promote *applied* critical thinking behaviors, to think critically and evaluate a technology requires a person to "step out" of this situated activity [2] in order to reflect on and evaluate the task. Collaborative learning can be used to achieve this by supporting reflection, by enabling students to take others' perspectives, discuss ideas, and come to a common understanding [32,37].

Exploration of sensor-based technologies

Sensors are commonly used as a part of a toolkit in learning contexts. For example, widely used physical computing toolkits like Arduino [40], micro:bit [6], and SAM Labs [33] provide engaging ways of experimenting, programming and making with sensors. However, research has not investigated how their properties can be appropriated to also teach critical reflection about the sensor and act of sensing itself.

Simultaneously, a number of studies have investigated how different sensors can be used to enable learners to reflect on the insights that sensor data can provide about themselves

and the environment. For instance, the Ambient Wood system showed how sensor-based systems can enhance curiosity and reflection when learning about the unseen data outdoors [29]. More recently, PhysiKit investigated how physical ambient visualizations could bring environmental sensor data to life in the family home and promote shared discussion about sensor technologies [14]. The ThinkActive project also explicated the opportunities and barriers to using wearable sensors in schools to enable children to reflect on their physical activity [10]. Others have investigated the use of personal sensors as a way of teaching children about statistical concepts like means and outliers, with much success, by capitalizing on embodied interaction with the sensors to enable children to tie these abstract concepts to their implicit knowledge of their bodies [19,20]. There is much scope in exploiting sensor properties – like their precision and ambiguity – to engender different types and levels of exploration of how they collect data [30]. While to date, much work has demonstrated how interacting with sensors can promote reflection about data, there is still a need to develop materials and curricula that can promote the range of critical thinking skills *per se*, for example about how sensor data is gathered and in what contexts it is accurate, reliable or informative. Next, we describe how we went about achieving this.

THE STUDIES

We designed a one-off, 90 minute session comprising open-ended, exploration-based activities for children, aged 9 to 11, that was aimed at engaging them in critical thinking about sensors, sensor data and the act of sensing. One of the aims of the session was to enable the children to learn how sensors can be both ambiguous and reliable to different extents. The focus was on developing a pedagogical framing that could support thinking about sensor accuracy (the extent to which the value detected by a sensor matches a true value), their reliability (how consistently accurate the data is), and how informative sensors are in different contexts.

An important part of supporting critical thinking is to determine how to provide appropriate guidance during the learning process, and how to encourage students to verbally reflect on learning activities with their peers and teachers. With this in mind, we provided exploratory tasks together with guidance structures, aimed to promote collaboration between peers, flexible support from instructors, and in situ reflection. Specifically, our pedagogical framing comprised the following 4 steps:

1. *Introduce sensors and sensing.* Verbally define *physical sensors and sensing* at the beginning of the session, with examples, to introduce the children to the concepts that they would be using during the exploratory activities.
2. *Frame the exploration of data collection in relation to the self and the environment.* Enable the children to engage in collecting and visualizing personal and environmental data using sensors in an open-ended way, together with providing suggestions for what to explore during the process.
3. *Encourage verbalization and reflection throughout:* Get the children to work in pairs/groups to enable collaborative learning to happen throughout the session, by providing multiple opportunities for them to show and tell, test their hypotheses and explain their discoveries to one another.
4. *Engage the children in a reflective discussion:* Prompt the children to reflect on their experiences with exploring the sensors, by engaging them in a reflective discussion supported by the instructor.

The Sensor Toolkit Used: The Magic Cubes

In steps 2 and 3 of the sessions, we used the Magic Cubes toolkit [18,39] to enable the children to explore and visualize sensor data (see Figure 2). The Magic Cubes are Arduino-based, wireless, hand-sized interactive cubes that contain a number of different embedded sensors. The cubes can be used to visualize sensor readings, and can be grasped, carried and shared easily, and in so doing encourage experimentation while exploring data [16,18].

Figure 2: The Magic Cubes toolkit comprises hand-sized cubes with embedded sensors and collocated numeric or symbolic visualizations.

Five different sensors were used with the cubes, for measuring the human body or environment (see Table 1). These were (i) a galvanic skin response (GSR) sensor; (ii) a pedometer, (iii) a pulse sensor; (iv) a temperature sensor and (v) a light sensor. For each, the cubes were programmed to provide a real-time data reading using a numeric (e.g., current temperature) or symbolic (e.g., real-time beating heart) display visualized on the embedded LED matrix. For the exploration-based component of the session, the children were given 7-10 minutes to explore each of the five sensors.

Field Journals to Scaffold Exploration and Reflection

To help the children explore the sensors and to promote reflection, we provided the children with *field journals* that they could flexibly use - a booklet of activity sheets that included suggestions for what to explore for each sensor, and questions to trigger reflection about the sensor properties and functionality. These were designed as a guide for the children's interactions in situ; for each sensor, they included three types of guidance: constrained tasks, open-ended tasks and reflective tasks.

Constrained Tasks. For each sensor, the field journal included a number of constrained tasks to help the children get started with exploring the sensor, and to reveal its complexities and ambiguities. For example, a GSR reading takes several seconds to change, which can be difficult to grasp when first interacting with it. The field journal asks the children to "take a deep, sharp breath in" and observe how long the sensor reading takes to change. This enables them to see a baseline of how rapidly and how much the reading changes. For the pedometer, the journal asks the children to walk around, count the number of steps they take, and compare this to the number on the pedometer. This enables them to see how accurate and reliable the sensor is when walking normally.

Open-ended Tasks. The journals also included open-ended suggestions to encourage more creative exploration of the sensors. These were aimed at building on the children's knowledge of their bodies and the environment when using the sensors. For example, the temperature sensor activity asks the children to get the cubes to display the hottest and coldest temperature possible, using their knowledge of temperature differences between materials and areas of the classroom. The journals also ask the children to try "tricking" the sensors. For example, the pulse sensor activity requires figuring out how they can get the pulse sensor to think their heart is beating faster than it is. We reasoned that in order to be able to trick the sensor, the children would have to analyze how it works and apply that knowledge to a new context. Table 1 provides additional examples of the open-ended tasks that were suggested.

Reflective writing. For each sensor, the children were encouraged to write down what they found to be interesting about the sensors, whether the display showed the values they expected, and how they tried to "trick" the sensors - successfully or unsuccessfully. This was intended to promote reflection and discussion during the hands-on portion of the session – especially about what the sensor actually measures, how it does this, and why the sensor reading might be wrong.

Participants

The study took place in five different classrooms at three different mixed-gender, mainstream schools in England (see Table 2). The class sizes ranged between 12 and 24 children. Four of the sessions were held in Year 6 classes (with children aged 10-11) and one in a Year 5 class (with children aged 9-10). A total of 86 children participated in the study.

School/Class		Year	Age	Number
1		5	9-10	24
2		6	10-11	15
3	Class 1	6	10-11	18
	Class 2	6	10-11	17
	Class 3	6	10-11	12

Table 2: The participants in each session.

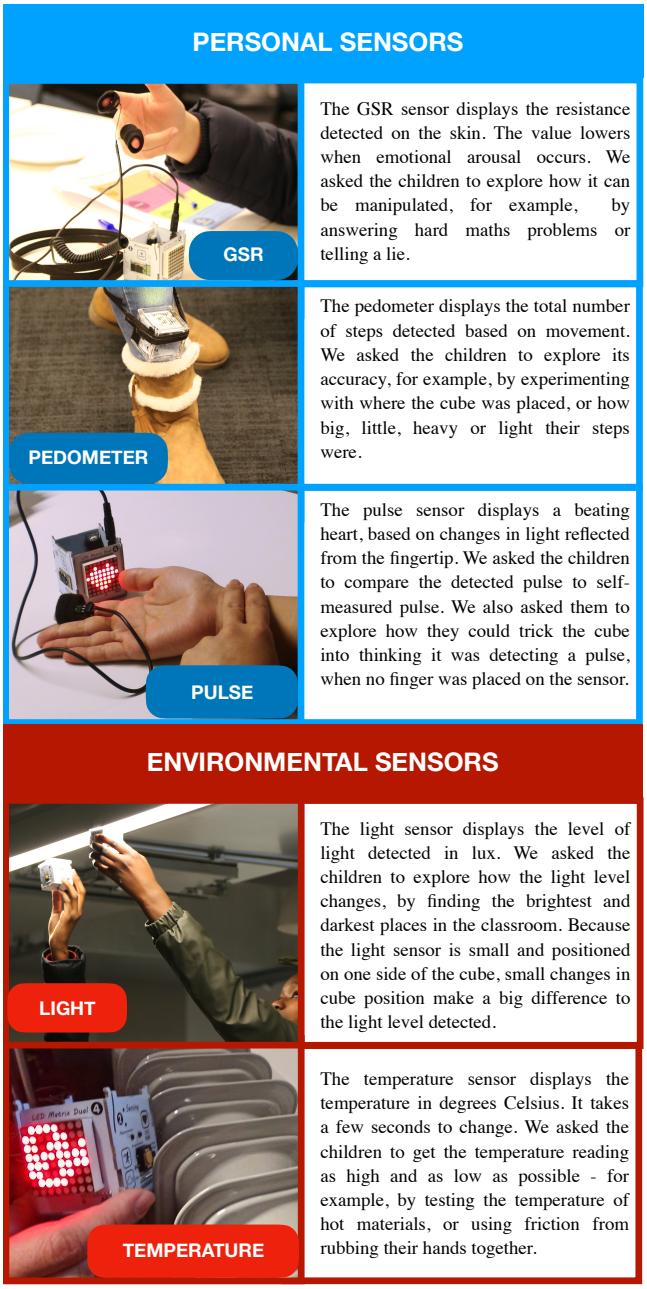

Table 1: The five sensors used for the study.

Procedure

The children were given a consent form for their parents to sign. On the day, the researcher began by informing the children of the purpose of the study and asked the children for their consent to be video and audio recorded; all agreed. In each session, the children were asked to work in pairs so as to promote dialogue and collaborative learning. They chose their own partners. For each session, a teacher from the school was present, as well as the researcher, and up to two additional research assistants (depending on their availability and the class size). The role of the researcher and the research assistants was to facilitate the sessions and guide the children through the tasks. The teachers were also

encouraged to take an active role in facilitating the session. The session started with a discussion with the class about what sensors and sensing devices are, and how they relate to everyday life. The children next spent a total of ~50 minutes exploring the five sensors in their pairs. This was followed by a class discussion, led by the researcher, where the children were asked to reflect on what they had discovered about the sensors, and to abstract away from the hands-on task to discuss the accuracy, reliability and informativeness of sensors in general.

Data Collection

Video cameras and audio recorders were placed throughout the classroom. This was done to gather continuous audiovisual data of the children's interactions and dialogue when carrying out the tasks to be recorded, and to capture their individual conversations. The audio recorders were placed on each desk; the video cameras were distributed so as to record both close shots of children sitting at their desks, and an overview of the classroom that captured the instructors (i.e., the teacher, researcher and research assistants) and the children's interactions when not at their desks. The children were also asked to use and fill out their field journals during the session.

Data Analysis

The recordings of the children's dialogue from the audio recorders were transcribed, and matched to corresponding video recordings of them interacting with the Magic Cubes. A qualitative approach of analyzing meaningful events that related to envisioned critical thinking outcomes was used when analyzing the children's physical interactions and dialogue, with both their peers and the instructors. Specifically, first, the video and audio recordings were iteratively viewed and annotated (total ~1100 minutes of footage across all the cameras) to create content logs of the sessions.

Next, a classification system of envisioned critical thinking outcomes was used to guide the analysis. This was based on the description of critical thinking introduced in the background section. Table 3 shows the classification system used, describing the mapping of the three envisioned critical thinking outcomes with a description of the outcome and a question that was used to guide the analysis. Based on this, "meaningful events" were identified in the dataset that evidenced the envisioned critical thinking outcomes, for example, a pair of children observing a counterintuitive property of a sensor, and using this to reflect on what the sensor measured and how. The identified meaningful events were transcribed in terms of the dialogue between the children, together with annotations about what the children were exploring at the time of the event, and how they were interacting with both the Magic Cubes, the field journals and with others around the classroom. Next, our findings are structured in terms of the three envisioned outcomes. They provide a qualitative account of when, to what extent and in what conditions the three different outcomes related to critical thinking were found to occur during the sessions.

Envisioned Critical Thinking Outcome	**Description of Outcome**	**Question Driving Analysis**
1. Extrapolate **what** the sensor actually measures and **how**	Although some sensors measure what their name indicates, others do not. For example, a GSR sensor measures skin resistance. Perceiving this relationship is viewed as corresponding to understanding what the sensor measures and how.	Does exploration enable children to understand what different sensors measure and be able to describe relationships between actions (e.g., telling a lie) and the sensor reading (e.g., the GSR value decreasing)?
2. Reason about **why** and **when** the sensor may not be working as expected	Reasoning about why something is not working as expected requires critical thinking processes including applying understanding, analyzing evidence, and making inferences. For example, applying the knowledge that a GSR sensor measures moisture underpins the inference that if the sensor is wet, the reading is likely not accurate.	Does exploration enable children to analyze and infer when a sensor might become inaccurate, uninformative or unreliable? What support mechanisms, if any, are required for them to engage with this form of reasoning?
3. Reason about the **accuracy, reliability** and **limitations** of sensors **in general**	It is one thing to analyze how a specific sensor works in practice, but another to extrapolate this when reflecting on sensors in general.	Do pairs of children explicitly discuss and evaluate factors that impact sensors' accuracy, reliability and informativeness overall? Do they do this when discovering the sensors, or is instructor facilitation required to promote explicit discussion?

Table 3: The classification system used in the analysis, based on the three envisioned critical thinking outcomes for this study.

FINDINGS

Overall, the findings showed that the children were able to move between understanding how the sensors worked and reflecting on their properties, while carrying out the open-ended exploratory tasks, where they collected personal data and came up with ways of testing hypotheses about how accurate and reliable the data was. This was evidenced by many instances of them verbally reflecting about data values that were unexpected, and of hypothesizing why the sensor data was not always accurate or reliable. Many playful, creative and collaborative interactions were observed taking place amongst the pairs. The children were found to be able to engage in aspects of critical thinking throughout the exploratory tasks. However, it was more difficult for them to engage in generalizing from their specific experiences to discuss the accuracy and reliability of sensors overall; to do so, they needed to be prompted by the teacher/researchers. In some ways, this is to be expected, given that it is something they are not used to talking about. In sum, the hands-on approach adopted here was effective at encouraging the children to begin to engage in aspects of critical thinking, which we examine next in terms of the envisioned outcomes.

Envisioned outcome 1: extrapolate *what* the sensor actually measures and *how*

The first question addressed was whether and how exploring the data gathered by the sensors would enable the children to understand and describe what the sensor measures and how it does this. When starting to interact with each sensor, the children were faced with the challenges of localizing the sensors on the cubes, figuring out how to position them on the body, and understanding what the values and symbols on the LED matrix meant. It was observed during all of the sessions, that when receiving a new sensor to experiment with, a majority of the children dived into exploring how it worked without first trying the suggestions provided in the field journal. Instead, the children flexibly mixed experimenting with the sensors in their pairs, with utilizing the variety of support structures available around the classroom – including the support of the instructors (the research team and the teachers) and the field journals.

For example, for the sensors that were externally attached to the cubes with a wire – that is, the GSR and the pulse – the finger gloves made it evident where the sensors were located. However, in order to learn how to use them, the children first had to figure out how to place their fingers inside the gloves to elicit an accurate sensor reading, specifically by placing their fingertips directly on the electronic components, and experimenting with how much pressure to put on the sensors. For the GSR sensor, some of the children were observed placing the electrodes on their fingernails rather than on their fingertips. This meant that the sensor would not be able to measure the change in resistance from sweat gland activity. In these instances, they asked the instructors for help, who explained to the children how to correctly place electrodes on their fingertips to enable them to use it correctly. Other times, when they did not understand how to use a specific sensor they reverted to looking at the instructions in the field journals, for example:

> C1: Does it go on your middle finger?
> C2: Read what it says on the sheet!
> C1: Um ok – [reads] 'Hint: keep your finger on the top of the green LED light, you might have to…' LED light.. Oh, that LED light!

Another challenge that the children faced when learning about what the sensors measure and how was to interpret the data visualized on the LED matrix of the cubes. It was assumed that this would be more difficult for the three personal sensors used – the GSR, pulse and the pedometer – all of which measure indirect indicators of a phenomenon, rather than the phenomenon itself. Specifically, the GSR sensor measures the resistance of the skin as an indicator of emotional arousal; the pulse sensor measures the amount of light reflected on the fingertip as an indicator that the heart has beaten; and the pedometer measures whether the movement of the cube itself is in a range likely to indicate that a step was taken, assuming the cube is strapped to the body.

One of the ways in which the children were found to engage in reflecting about these properties was through answering impromptu questions raised by the instructors. For example, the researcher (R) noticed that a pair who had said they were done with the pulse sensor had not filled out the section in the field journal that was about tricking the sensor. The researcher prompted the pair to engage with the section:

> R: Have you tried tricking it yet?
> C1: How do you trick it?
> R: So um, you have to figure out when it doesn't work. [...] it's not on your finger and it's still kind of giving a heartbeat, right?
> C1: Yeah
> R: Why do you think that is?
> C1: (confidently) The table. Cause we're like jiggling the table – and going like that --
> R: -- Yeah! So what do you think it's actually measuring?
> C1: Like movement?

Here, the researcher explained what "tricking" the sensor meant, and next asked the pair to make a hypothesis as to why the LED matrix of the cube indicated that a pulse has been detected, when the fingertip was not on the sensor. In this example, C1 was incorrect in saying that the pulse sensor is measuring movement; however, this instance led the pair to start hypothesizing about other ways to "trick" the sensor. They then began experimenting with the sensor in other ways, such as tapping it, which also led the sensor to detect a false 'pulse'. The pair later participated in the classroom discussion, where they were able to correctly hypothesise that the pulse sensor reflects light.

With the exception of the GSR sensor, the children were seen to spend very little time reflecting on how the visualizations mapped onto the phenomena being measured. The way the data was represented seemed to be easy to understand – for example, the numerical light level represented on the LED matrix increased in brighter places, and the LED matrix flashed a heart when a heartbeat was detected with the pulse sensor. However, for the GSR sensor, most of the children found the directionality of the change in the values confusing. This was because the sensor measures resistance, a value that *decreases* with emotional arousal (e.g., stress when telling a lie) – which is counterintuitive, if assuming that telling a lie makes the sensor value rise. Because of this confusion about what an increase versus a decrease in the GSR reading meant, the children spent much time trying to test it – for example, by repeatedly asking whether each other the GSR value goes up or down when telling a lie.

Moreover, when some of the children placed the GSR sensor on their fingertips, the value was as low as 0 or 1. This happened when they had wet fingertips, or when the sensor was wet from someone who had used it before. In these instances, there was no room for the sensor value to drop further, which impeded exploration of the data. However, it was found that experiencing this phenomenon sometimes had the positive effect of enabling the children to reason about how the GSR sensor might work and what it might measure. For example, one child reflected, *"I asked everyone everything, and I got 0!"*. After being asked why this happened, he replied that, *"it was wet when I put it on!"*, suggesting that he had reasoned that moisture played a part in the GSR data.

For the pedometer, it was found that through the exploratory task, the children were able to make a distinction between the measurement of movement and the measurement of the more abstract concepts of steps – and moreover, reflect on why this mattered in context of accuracy. For example, while having the pedometer strapped to her wrist, a student noticed that it was adding steps when she moved her hand, and later inferred how an everyday pedometer might work in practice:

> You see like, when you wear those thingies like – the Fitbits and stuff – it's on your wrist. […] I think it's like checking like when you move your hands around. I think it's going to the rhythm of that, not actually [your step]."

Envisioned outcome 2: reason about *why* and *when* the sensor may not be working as expected

It was found that the children were, to a large extent, able to discover that the sensors were not always reliable or accurate. For example, they were able to observe that the pulse sensor was prone to inaccuracy, especially when the wire was moved, or when a finger was placed too lightly on the sensor. They were also able to observe that how informative a sensor was varied depending on the context of use – for instance, that the GSR sensor is informative as a way of measuring changes in emotional arousal, but not as a lie detector *per se*. However, how readily they engaged with these outcomes varied between sensors. Next, we describe how the types of interaction and observable effects afforded by the different sensors, enabled or impeded this.

Reflecting through embodied interaction

When exploring the pedometer, the children were seen to shake the cube, dance with the cube, jump around or walk without moving their arms. In doing this, they began to observe and reflect upon how the position in which they placed the cube on their body, as well as the type of movement they enacted, influenced the accuracy of the step count. For example, after attaching the pedometer cube to her wrist, and walking without moving her hands, a girl said:

> Let's say the pedometer was on my wrist, and over there [points to a narrow space between two desks], when I tried to get through it I couldn't move my hand back […] and I think when I moved my hand it counted the steps… And I didn't move my hand so it didn't count that as steps.

Reflecting on unexpected sensor behaviors

Another way the children reflected on sensor accuracy, reliability and informativeness was through observing unexpected sensor behaviors. This happened most frequently with the GSR sensor. Across all sessions and pairs, the children's favorite use case for the GSR was found to be using it as a lie detector, and they were observed to test out the GSR's lie-detecting capabilities in a diversity of ways. Specifically, they spent time asking each other playful questions and guessing if they were lying by checking if the GSR value displayed on the LED matrix of the cube changed. When using the GSR in this way, many of the children initially assumed that the sensor would be able to tell when someone was lying in all instances. However, experimenting with telling different types of lies and truths while wearing the GSR sensor enabled them to observe that the sensor was not consistently able to catch them when they were lying. Many types of questions were asked, including fairly innocuous ones (e.g., *"do you like chocolate?"*, *"have you ever teleported?"*) and more stressful ones (e.g., *"do you have a crush on someone in this class?")*. To their delight, they found that asking different kinds of questions triggered different levels of emotional arousal in the one answering, which were not always tied to lying or telling the truth. Sometimes, answering a question caused the GSR value to fall to as low as 1 or 0, while other times it stayed the same or increased slightly. For example, when one of the children lied about having teleported, the GSR value neither decreased nor increased, which, under the assumption that the value would drop when a lie was told, would indicate that the child had indeed teleported. In another example, one girl replied yes when she was asked if she had a crush on someone at school. Because this was stressful, the GSR value dropped quickly from a value of 140 to 47, prompting her partner to accuse her of lying, rather than because she was stressed that she was telling the truth:

C1: Do you actually? [pause; watching the sensor value, which decreases] You're lying!
C2: I'm not. It just went to 47… but I'm not!

These unexpected sensor responses were able to challenge the children's assumption of the GSR as an accurate lie detector, as well as enable them to question how informative the GSR sensor was when used in this way. For example, one of the children asked the instructor, *"what if you don't get stressed when you're asked a question? People don't always get stressed!"*

It was found that reflecting on unexpected sensor behaviors conversely also helped the children reflect on the first envisioned outcome – that is, understanding what the sensor measures and how. For example, a pair had been trying for several minutes to elicit a sensor response by asking each other to tell white lies (for example, by asking "*do you like pizza?*"), but noticed that the GSR sensor was not changing as they had expected. The class teacher stepped in at this point to explain the relationship between stress and moisture on the skin, as opposed to lying and moisture on the skin:

T: You know what? Why I don't think it works with that as much is because you're just saying a lie but you're not really feeling that stressed, whereas the reason it's doing it is because it's measuring moisture. But actually if somebody asked you something and you were quite under pressure and you had to lie, you'd feel more stressed than if you were telling the truth. Do you see what I mean?
C1: Yeah. Ok! What's a question she can get stressed on though?
T: What Kira, you don't like Harry Potter? [in a shocked voice]
C1: [sensor reading drops] 227! … It's getting higher then lower, then lower and then higher. You're at 200… 257. [pause] Ok. Are you scared? Of me?
C2: No [laughs]
C1: It went down, you are scared of me!

These types of unexpected sensor behaviours were also seen to trigger much reflection about the pulse sensor and the pedometer - for example, observing the inaccuracy of the pulse sensor when the heart animation flashed when moving the finger (e.g., "*every time I put my finger on it just flashes*"), or how the pedometer added a number of extra steps when it was picked up off the floor (e.g., "*I took it up with me to the table, and the number went up to 58!*").

In contrast, much less reflection and verbal reasoning was found to occur when using the light and temperature sensors. Instead, the children spent more time reasoning about the material properties of objects, for example, discussing why a rubber spatula is warmer than a metal table leg, or why pointing a cube towards the indoor light triggers a lower value than pointing a cube towards the sun. There were no observed instances of them reasoning about the sensor data itself. The lack of explicit reasoning about the sensor properties may have been because the light and temperature sensors are relatively easy to understand and use—that is, while they are not always accurate, they did not present any obvious unexpected behaviors that could be tied back to observed or experienced phenomena. For example, a light value in lux is difficult to relate to an exact light level in the real world, as is a temperature value. This afforded focusing mostly on what was to be measured, rather than the device used for measuring.

Envisioned outcome 3: reason about the accuracy, reliability and limitations of sensors *in general*

Compared to the first two envisioned outcomes, there was less evidence of the children talking about the accuracy and reliability of the sensors during the exploratory activities. However, there was more evidence of this happening when they were explicitly asked about them during the reflective discussion phase that followed the exploratory tasks, for example:

R: So, what does that tell you about sensors? Are they accurate?
C1: They're very accurate
C2: They're not very accurate [shaking head].
R: So, what does it depend on?
C3: They're accurate, but it's easy to trick them so you have to be careful how you use them. So, if you're like – if you're going too fast, then it won't detect it, if you're moving your legs too fast, it won't count the right amount of steps so you have to be careful how you actually use them.

Here, C3 builds on her classmates' responses, by relating the question about accuracy with her previous experiences from the exploration-based activity to support her point. She describes instances that she observed of the pedometer not working, to motivate her conclusion on the accuracy of the sensors. This suggests that the children were able to build an implicit knowledge of the limitations of the sensors through the hands-on activities—for example that their accuracy is dependent on how they are used—which they were able to reflect upon subsequently.

However, during the discussion phase, the children's responses did not always convey a *complete* understanding of the topics. For example, in one session, a pair of children mentioned that even when they did not have their fingers on the GSR sensor, it displayed a sensor reading. They discussed what this might mean: "*they're not quite accurate because when we took it out, there was nothing on the thing [sensor] – we didn't put our fingers in it, and it just changed the numbers*." This then triggered a discussion, where the researcher explained that the sensor has no way of knowing whether or not someone has placed a finger on the sensor, and instead constantly measures resistance, which is not necessarily telling of its **accuracy** but rather of how **informative** it is in a particular context. In sum, while the children were able to reason about the high-level topics by drawing on their experiences with the Magic Cubes, the

analysis suggests their understanding of the topics was not always complete.

DISCUSSION

The findings from the studies showed that the children were able to engage in critical thinking to a certain extent when reasoning about the data that they collected about their bodies and their environment using the Magic Cubes. In particular, there was much evidence that they understood that some sensors are not always accurate and do not always reliably measure phenomena that they are assumed to. While the children did not spontaneously talk about the data concepts they were introduced to – reliability, accuracy and how informative sensor data is – during the hands-on tasks, they were able to reflect on them afterwards, when prompted. Not taking a sensor reading at face value and wondering how it can vary depending on what someone does with a sensor was an important lesson that enabled the children to think more critically, for example, about what it means to measure GSR, and in what contexts it can be relied upon.

In answer to the research question about how exploration-based physical toolkits can enable critical thinking, we found that two properties of the exploratory activities designed for the Magic Cubes, in particular, were important for supporting reasoning and critical reflection. These were the provision of sensors that enabled the children to relate abstract data to a lived experience, and observable unexpected sensor behaviors that caused the children to "step out" of the hands-on activity to reflect about the data in relation to their actions.

Relating data to a lived experience

Promoting learning that capitalized on their awareness of their bodies and environment (see [15,19]) was found to help the children make connections between the sensor, the data collected and how it mapped onto the underlying activity that was being measured (e.g., moving, breathing, answering an embarrassing question). Enabling the children to explore their personal data – such as GSR, step count and pulse – together with concrete and easy to understand visualizations of the sensor values, was found to be another way of facilitating critical thinking. This also enabled them to think about what the reasons were when the data that was displayed was perceived to be inaccurate.

In contrast, they did not appear to notice values that might have been inaccurate or unreliable for the light and temperature sensors they collected. This suggests that it was more difficult to spot when a reading from one of these environmental sensors was wrong. One reason for this is it was not possible to establish a ground truth for these sensors in the same way as could be done for the personal sensors. While it is straightforward to relate an increasing value of light or temperature to a brighter or warmer place, it is harder to establish the accuracy of specific values in degrees Celsius or light level in lux without using another measuring device. This contrasts to establishing a ground truth of perceived stress, pulse rate or number of steps taken by capitalizing on embodied knowledge. Together, this demonstrates that enabling the child to manipulate what is being sensed on their bodies provides a personal testing ground that can foster the development of critical thinking skills.

Unexpected sensor behaviors

Another way that critical thinking was supported during the exploration process with the Magic Cubes was through experiencing unexpected sensor behaviors. The properties of the sensors that were used meant that they sometimes worked in ways that were ambiguous or counterintuitive. For example, the GSR value went down with stress level, instead of up; the pulse sensor reading was sensitive to changes in light; and the pedometer added steps on when the cube was dropped or shaken. Because these effects were readily observable, they promoted much verbal reflection between the children about how the sensors worked, and about when they broke their expectations. This suggests that a good strategy for promoting critical thinking is to provide activities which are meaningful to the child, and where the data collected with a sensor can at times be puzzling or be ambiguous (see [30]). This makes them stop and think why it is showing a given reading, especially if it is contrary to what they expect.

Components of critical thinking supported during exploration-based learning

Our analysis was framed in terms of three envisioned critical thinking outcomes: extrapolating what the sensor measures and how; reasoning about why and when the sensor may not be working as expected; and reasoning about the accuracy, reliability and limitations of sensors in general. Our findings lend much support for the ability of exploration-based learning to engage the children with the first two envisioned outcomes. It was found that, during the exploration-based task, the children were able to reason about the sensors while applying their understanding of how they work, experimenting with them and analyzing the data readings that they obtained using the Magic Cubes. These activities supported them in understanding what the sensor measures and how, which they then used to reason about why and when the sensors may not be working as expected. The cognitive processes that led to these outcomes were seen to feed into each other in both directions. The children often applied their understanding of how the sensors work to infer why they were working in unexpected ways. For example, some children were able to apply their understanding of the fact that the pedometer measures how much the cube has moved, to reason why it did not add steps when walking without moving their hands, if the cube was placed on their wrist. Conversely, by analyzing why the sensors were not working as expected, they refined their understanding of how the sensors worked. For example, observing that the GSR sensor reading did not change as expected when telling an innocuous lie, led some to infer that it was measuring values related to a stress response.

Thus, even though they were not explicitly asked to engage in a structured scientific enquiry process during the

exploratory activities, they did so, by way of making hypotheses, analyzing the observed data and inferring its meaning, to a larger extent than expected. This suggests that there in much promise for designing open-ended, hands-on activities when the goal is to promote curiosity and critical thinking about data. This is in line with other research on technology-mediated exploration of data for children, where promoting student-initiated exploration of phenomena with a technology has been found to enable scientific enquiry, even if this is not explicitly asked of the students [29,31].

However, despite the positive findings that the children engaged in critical thinking about *specific sensors* during the exploratory activities, the level of *abstract* critical thinking that they engaged in was limited. In particular, while they interacted with the Magic Cubes, no instances were found of them discussing, evaluating and judging *sensors in general* – our third envisioned critical thinking outcome. To explicitly evaluate and judge the limitations of sensors in general, they had to be probed by their teacher or the researchers. In some ways, this is to be expected, given the study was run as a one-off session in each school, and that the concepts of accuracy and reliability in the context of sensors were only introduced to the children at the start of the session. However, it suggests that there are limits to what can be achieved with open-ended learning activities, in particular, how children can abstract away from a specific hands-on task to relate it to more general principles. This supports previous literature on discovery learning – a sub-category of learning through exploration – which suggests that a level of cognitive guidance is important for enabling students to integrate the observations acquired from a hands-on, behavioral activity into more abstract patterns and principles [21].

Nevertheless, the way in which the children based their descriptions of how accurate, reliable or informative sensors are in general during the discussion session, was often by supporting their responses with what they had observed during the exploration process. While their understanding of the target concepts was not always complete, the hands-on experience had a positive effect on enabling them to evaluate and judge the reliability of the sensors and their ability to accurately sense certain phenomena. This suggests that the dovetailing of well-designed exploration-based activities and discussion during learning enables children of this age group to learn and reflect about abstract concepts such as those in IoT, providing the basic building blocks for more advanced critical thinking.

Support mechanisms for triggering critical thinking in a classroom context

It was found that a number of support mechanisms that were introduced in the study were effective in triggering critical thinking; these were a combination of working in pairs, using field journals, and instructor support. Turning to one of these forms of help, was most marked when a pair got stuck or observed an unexpected sensor effect (such as the GSR sensor not detecting a lie). Here, we observed them talking to each other about what to do next, checking the journals for guidance, or calling on the support of an instructor – all of which provided opportunities to verbalize and reflect on their experiences.

Finally, it was found that the exploratory activities often led to highly visible, loudly spoken and performative interactions. As noted, the interactions were often playful and, in some cases, competitive. Examples included children exclaiming in surprise when unexpected sensor responses were observed and dancing around the classroom. This type of highly charged and visible interaction concurs with previous research that suggests such performative acts can facilitate collaboration and communication [13]. Here, they helped the instructors monitor the activity in the class, and intervene at appropriate points when necessary. They also attracted the children to turn their attention to observe others around the classroom, and in this way promoted peer learning – as the children were able to monitor each other's actions and help other pairs when they noticed them struggling. The combination of learning activity and learning environment was thus effective at supporting practicing critical thinking. This suggests it is helpful to have flexible scaffolding in place when designing for exploration-based learning that is aimed at teaching children to reason about computing concepts at a deeper level. Here, having the choice of asking others, observing others, having an instructor-led discussion or looking up suggestions in the field journals, provided a number of mechanisms for this.

CONCLUSION

This study has shown how it is possible to encourage children to begin to understand that sensing isn't just about reading off data from a device; depending on how the sensor is used and in what context, sensor data can be inaccurate, unreliable or uninformative. This in turn means that sometimes the data from sensors can be relied upon, but other times that is not the case. Understanding the basic principles of accuracy and reliability are important stepping stones for learning about other topics, for example, how to filter noise and capture patterns in datasets, and thinking critically about how the data that makes up a dataset can influence bias in IoT, AI and other paradigms. Our study has demonstrated how to embed the process of critical thinking in learning about computing in such a way that enables children to readily and enjoyably engage with these topics when beginning to learn about computing. As such, it can better equip them with not just the ability to understand how an aspect of a technology works, but also the ability to question it and probe its limitations.

ACKNOWLEDGMENTS

This research has been made possible by the EPSRC and the BBC through an iCASE scholarship (Award Number 1623937). We would like to give a profound thanks to the teachers and children who participated in the studies for their time and enthusiasm. We would also like to thank the UCL

students who assisted with the classroom sessions for their valuable support.

SELECTION AND PARTICIPATION OF CHILDREN

To recruit participating classrooms, the Magic Cubes were showcased at festivals and events related to computing education in schools. At these events, contact information was collected from teachers who expressed an interest in trialing the Magic Cubes in their classrooms, which we then followed up on. Prior to the study, parents of children in the participating classrooms were provided with information sheets outlining the aims of the study, the data that would be collected and how this data would be used. All children who participated had written parental consent. On the day of the study, the participating children were provided with the same information in simplified language, both in written and verbal form, given the opportunity to ask questions, and asked to fill out consent forms.

REFERENCES

[1] Philip C. Abrami, Robert M. Bernard, Evgueni Borokhovski, Anne Wade, Michael A. Surkes, Rana Tamim, and Dai Zhang. 2008. Instructional interventions affecting critical thinking skills and dispositions: A stage 1 meta-analysis. *Review of Educational Research* 78, 4: 1102–1134.

[2] Edith K. Ackermann. 1996. Perspective-Taking and Object Construction: Two Keys to Learning. In *Constuctionism in Practice: Designing, Thinking, and Learning in a Digital World (Kafai, Y., and Resnick, M., Eds.)*. Mahwah, New Jersey: Lawrence Erlbaum Associates. Part 1, Chapter 2., 25–37.

[3] Alissa N. Antle and Alyssa F. Wise. 2013. Getting Down to Details: Using Theories of Cognition and Learning to Inform Tangible User Interface Design. *Interacting with Computers* 25, 1: 1–20. https://doi.org/10.1093/iwc/iws007

[4] Sharon Bailin, Roland Case, Jerrold R. Coombs, and Leroi B. Daniels. 1999. Common misconceptions of critical thinking. *Journal of curriculum studies* 31, 3: 269–283.

[5] Sharon Bailin and Harvey Siegel. 2002. Critical thinking. *The Blackwell guide to the philosophy of education*: 181–193.

[6] BBC micro:bit. 2016. Micro:bit: get creative, get connected, get coding. Retrieved from https://www.microbit.co.uk

[7] Karen Brennan and Mitchel Resnick. 2012. New frameworks for studying and assessing the development of computational thinking. In *Proceedings of the 2012 annual meeting of the American Educational Research Association, Vancouver, Canada*, 25.

[8] Robert H. Ennis. 1985. A logical basis for measuring critical thinking skills. *Educational leadership* 43, 2: 44–48.

[9] Peter Facione. 1990. Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction (The Delphi Report).

[10] Andrew Garbett, David Chatting, Gerard Wilkinson, Clement Lee, and Ahmed Kharrufa. 2018. ThinkActive: Designing for Pseudonymous Activity Tracking in the Classroom. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (CHI '18), 7:1–7:13. https://doi.org/10.1145/3173574.3173581

[11] D. Randy Garrison. 1992. Critical thinking and self-directed learning in adult education: An analysis of responsibility and control issues. *Adult education quarterly* 42, 3: 136–148.

[12] Diane F. Halpern. 1992. A Cognitive Approach to Improving Thinking Skills in the Sciences and Mathematics. *Enhancing thinking skills in the sciences and mathematics*.

[13] Eva Hornecker and Jacob Buur. 2006. Getting a Grip on Tangible Interaction: A Framework on Physical Space and Social Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '06), 437–446. https://doi.org/10.1145/1124772.1124838

[14] Steven Houben, Connie Golsteijn, Sarah Gallacher, Rose Johnson, Saskia Bakker, Nicolai Marquardt, Licia Capra, and Yvonne Rogers. 2016. Physikit: Data Engagement Through Physical Ambient Visualizations in the Home. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (CHI '16), 1608–1619. https://doi.org/10.1145/2858036.2858059

[15] Robert JK Jacob, Audrey Girouard, Leanne M. Hirshfield, Michael S. Horn, Orit Shaer, Erin Treacy Solovey, and Jamie Zigelbaum. 2008. Reality-based interaction: a framework for post-WIMP interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 201–210.

[16] Rose Johnson, Venus Shum, Yvonne Rogers, and Nicolai Marquardt. 2016. Make or Shake: An Empirical Study of the Value of Making in Learning About Computing Technology. In *Proceedings of the The 15th International Conference on Interaction Design and Children* (IDC '16), 440–451. https://doi.org/10.1145/2930674.2930691

[17] Emily R. Lai. 2011. Critical thinking: A literature review. *Pearson's Research Reports* 6: 40–41.

[18] Zuzanna Lechelt, Yvonne Rogers, Nicola Yuill, Lena Nagl, Grazia Ragone, and Nicolai Marquardt. 2018. Inclusive Computing in Special Needs Classrooms: Designing for All. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 517.

[19] Victor R. Lee and Joel Drake. 2013. Quantified recess: design of an activity for elementary students involving analyses of their own movement data. In *Proceedings of the 12th international conference on interaction design and children*, 273–276.

[20] Victor R. Lee, Joel R. Drake, and Jeffrey L. Thayne. 2016. Appropriating quantified self technologies to support elementary statistical teaching and learning.

IEEE Transactions on Learning Technologies 9, 4: 354–365.

[21] Richard E. Mayer. 2004. Should there be a three-strikes rule against pure discovery learning? *American psychologist* 59, 1: 14.

[22] D. R. Newman, Chris Johnson, Brian Webb, and Clive Cochrane. 1997. Evaluating the quality of learning in computer supported co-operative learning. *Journal of the American Society for Information science* 48, 6: 484–495.

[23] Sofia Papavlasopoulou, Michail N. Giannakos, and Letizia Jaccheri. 2017. Empirical studies on the Maker Movement, a promising approach to learning: A literature review. *Entertainment Computing* 18: 57–78.

[24] Partnership for 21st Century Learning. 21st Century Learning for Early Childhood Framework. Retrieved May 24, 2019 from http://www.p21.org/our-work/elf

[25] Jean Piaget. 2013. *The construction of reality in the child*. Routledge.

[26] Sara Price and Yvonne Rogers. 2004. Let's get physical: The learning benefits of interacting in digitally augmented physical spaces. *Computers & Education* 43, 1–2: 137–151.

[27] Mitchel Resnick, John Maloney, Andrés Monroy-Hernández, Natalie Rusk, Evelyn Eastmond, Karen Brennan, Amon Millner, Eric Rosenbaum, Jay Silver, and Brian Silverman. 2009. Scratch: programming for all. *Communications of the ACM* 52, 11: 60–67.

[28] Jennifer A. Rode, Anne Weibert, Andrea Marshall, Konstantin Aal, Thomas von Rekowski, Houda El Mimouni, and Jennifer Booker. 2015. From computational thinking to computational making. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 239–250.

[29] Y. Rogers, S. Price, G. Fitzpatrick, R. Fleck, E. Harris, H. Smith, C. Randell, H. Muller, C. O'Malley, D. Stanton, M. Thompson, and M. Weal. 2004. Ambient Wood: Designing New Forms of Digital Augmentation for Learning Outdoors. In *Proceedings of the 2004 Conference on Interaction Design and Children: Building a Community* (IDC '04), 3–10. https://doi.org/10.1145/1017833.1017834

[30] Yvonne Rogers and Henk Muller. 2006. A framework for designing sensor-based interactions to promote exploration and reflection in play. *International Journal of Human-Computer Studies* 64, 1: 1–14.

[31] Yvonne Rogers, Sara Price, Eric Harris, Ted Phelps, Mia Underwood, Danielle Wilde, Henk Muller, Cliff Randell, Danae Stanton, and Helen Neale. 2002. Learning through digitally-augmented physical experiences: Reflections on the ambient wood project.

[32] Jeremy Roschelle and Stephanie D. Teasley. 1995. The Construction of Shared Knowledge in Collaborative Problem Solving. In *Computer Supported Collaborative Learning*, Claire O'Malley (ed.). Springer Berlin Heidelberg, 69–97. https://doi.org/10.1007/978-3-642-85098-1_5

[33] SAM Labs. Sam Labs: Discover the fun in STEAM and coding in your classroom. Retrieved September 1, 2019 from https://samlabs.com

[34] Sue Sentance, Jane Waite, Steve Hodges, Emily MacLeod, and Lucy Yeomans. 2017. Creating Cool Stuff: Pupils' Experience of the BBC micro: bit. In *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education*, 531–536.

[35] Sue Sentance, Jane Waite, Lucy Yeomans, and Emily MacLeod. 2017. Teaching with physical computing devices: the BBC micro: bit initiative. In *Proceedings of the 12th Workshop on Primary and Secondary Computing Education*, 87–96.

[36] UK Department of Education. 2016. National curriculum in England: design and technology programmes of study - GOV.UK. Retrieved from https://www.gov.uk/government/publications/national-curriculum-in-england-design-and-technology-programmes-of-study/national-curriculum-in-england-design-and-technology-programmes-of-study

[37] Lev Semenovich Vygotsky. 1978. *Mind in society: the development of higher psychological processes*. Harvard University Press.

[38] CS4ALL Blueprint Beta. *CS4All Blueprint - Beta*. Retrieved January 12, 2020 from https://blueprint.cs4all.nyc/

[39] MagicCubes – MagicCubes: the cool way to experiment with sensors, lights and coding. Retrieved January 6, 2020 from http://codeme.io/wordpress/

[40] Arduino - Home. Retrieved January 6, 2020 from https://www.arduino.cc/