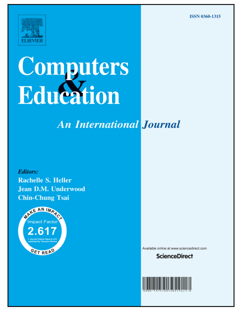
Accepted Manuscript



A systematic review on teaching and learning robotics content knowledge in K-12

Liying Xia, Baichang Zhong

PII: S0360-1315(18)30243-4

DOI: 10.1016/j.compedu.2018.09.007

Reference: CAE 3450

To appear in: *Computers & Education*

Received Date: 8 March 2018

Revised Date: 6 September 2018

Accepted Date: 15 September 2018

Please cite this article as: Xia L. & Zhong B., A systematic review on teaching and learning robotics content knowledge in K-12, *Computers & Education* (2018), doi: https://doi.org/10.1016/j.compedu.2018.09.007.

A systematic review on teaching and learning robotics content knowledge in K-12

Liying Xia [a], **Baichang Zhong** [a *]

[a] School of Education Science, Nanjing Normal University, Jiangsu, China

[*] Corresponding author. School of Education Science, Nanjing Normal University, No.122, Ninghai Road, Nanjing, 210097, China.

E-mail address: zhongbc@163.com (B. Zhong).

A systematic review on teaching and learning robotics content knowledge in K-12

Abstract

This paper aims to review high-qualified empirical studies on teaching and learning robotics content knowledge in K-12 and explore future research perspectives of robotics education (RE) based on the reviewed papers. After a systematic search in online database via keyword search and snowballing approach, 22 SSCI journal papers are included in this review. Nine major factors are examined for each paper: sample groups, duration, robot types, robotics content knowledge, study type, intervention approaches, measurement instruments, major findings and instructional suggestions. The results indicate that: (1) most empirical studies were endured less than two months with a small sample size, the largest sample group was elementary school students, and most studies used LEGO robots; (2) more than half of the studies conducted a non-experimental research design, and observation, questionnaire, interview and evaluation of artifacts are commonly-used measurement instruments; and (3) instructional suggestions proposed in the 22 papers can be clustered into four themes: open environment, targeted design, appropriate pedagogy and timely support. Overall, the 22 papers suggest that RE shows great educational potential in K-12, however, there are indeed situations in which RE did not bring significant improvement in student learning. In view of this, we prospect the future research directions of RE and propose that more intervention studies with rigorous research design could be conducted in K-12.

Keywords

Elementary education; Secondary education; Improving classroom teaching; Pedagogical issues; Teaching/learning strategies

1. Introduction

Robotics education (RE), a promising way for getting students engaged in Science, Technology, Engineering, and Mathematics (STEM), has greatly attracted the interest of researchers and teachers alike, from kindergarten to university during the past years. According to Gomoll, Hmelo-Silver, Šabanović, and Francisco (2016) and Master, Cheryan, Moscatelli, and Meltzoff (2017), positive experience with RE would generate greater interest and self-efficacy among students, particularly for girls, thus inspiring them to engage in STEM in the future. A family of studies (Bers, Flannery, Kazakoff, & Sullivan, 2014; Chen, Shen, Barth-Cohen, Jiang, Huang, & Eltoukhy, 2017; Leonard et al., 2016) demonstrated that while engaging in construction-based robotics activities, students can be both interested in and able to learn many aspects of robotics, programming, and computational thinking. Several studies (Shih, Chen, Wang, & Chen, 2013; Sullivan, 2008) suggested that the utilization of robots in school settings could lead to the improvement of students' science literacy.

This increasing interest in RE calls for a systematic review on the previous studies for the purpose of acting as a "pathfinder" for future research. Thus, three representative literature reviews were found as follows. Benitti (2012) presented a literature review to identify the potential contribution of robotics as an educational tool. The results showed that robotics is useful to aid the understanding of concepts related to STEM, but there is a lack of research with quantitative assessment on learning performance. Mubin, Stevens, Shahid, Mahmud, and Dong (2013) found that robotics are primarily used to provide language, science or technology education, and that a robot could play the role of a tutor, tool or peer in the learning activity. Toh, Causo, Tzuo, Chen, and Yeo (2016) examined the use of robots in early childhood and lower level education. The result revealed that robot's influence on children's skills development could be grouped into four major categories: cognitive, conceptual, language and social (collaborative) skills.

These literature reviews, however, have several limitations. They: (1) did not use multiple search strategies to retrieve sufficient papers, since comprehensive searches are indispensable for a full-fledged review (Kitchenham, Brereton, Budgen, Turner, Bailey, & Linkman 2009); (2) did not ensure the quality of reviewed papers, as the selected papers included many conference papers; (3) did not focus on teaching and learning robotics content knowledge, which is an important component of STEM; and (4) did not put emphasis on the empirical studies in K-12, leading to a lack of understanding of the current state of RE in K-12.

Considering these limitations, the purpose of this study is twofold: (1) to systematically review high-quality empirical studies on teaching and learning robotics content knowledge in K-12, and (2) to explore future research perspectives of RE based on the reviewed papers. The term "empirical study", which means "knowledge based on real-world observations or experiment," is used here to describe field-based research which uses data gathered from naturally occurring situations or experiments, rather than via laboratory or simulation studies, where the researchers have more control over the events being studied (Flynn, Sakakibara, Schroeder, Bates, & Flynn, 1990). In order to systematically examine the empirical evidence on RE, this review is guided by the following questions.

Q1: How have robotics been incorporated into K-12?

Q2: What intervention approaches are effective in teaching and learning robotics content knowledge?

Q3: What implications for teaching are indicated by these empirical studies?

2. Methods

For a full-fledged review, search strategy is key to ensure a good starting point for the identification of reviewed papers and ultimately for the actual outcome of the review (Kitchenham et al., 2009). To select useful and high-quality papers, we performed a keyword search in international online bibliographic database ISI Web of Science. Due to the limited papers selected via keyword search, we used snowballing as a search approach to get as many papers as possible in October 2017 and updated in December 2017. Snowballing approach refers to using the reference list of a paper or the citations to the paper to identify additional papers for a literature review (Wohlin, 2014). The following three-step iterative process outlined by Kitchenham et al. (2009) and Wohlin (2014) was adopted in each round of our snowballing approach. Once it is difficult to find new papers in the iterations using both backward and forward snowballing, the loop can be ended.

Step 1. Start set: identify a starting set of papers for the snowballing approach from the leading journals in the field.

Step 2. Backward snowballing: go backward by looking at the reference lists of the papers identified in step 1 to determine which papers are qualified for your literature review.

Step 3. Forward snowballing: go forward by looking at the citations to the papers identified in step 1 to determine which papers are qualified for your literature review.

The following inclusion criteria were implemented to determine which papers could be included in this review.

Inclusion Criteria 1: Papers published in SSCI journals.
Inclusion Criteria 2: Papers belonging to empirical study.
Inclusion Criteria 3: Papers reported studies conducted in K-12 for students.
Inclusion Criteria 4: Papers involved in teaching and learning robotics content knowledge.

2.1. Keyword search

Initially, we used the search string "robot* AND (education OR instruction OR school OR curriculum OR curricula OR course OR student OR child* OR learn*)" to identify SSCI journal papers written in English in the ISI Web of Science. The results returned 344 papers. We quickly analyzed the titles and abstracts of these papers with reference to the above four inclusion criteria. As a result, 12 papers were left.

2.2. Snowballing approach

After finishing the keyword search, we employed the first-round snowballing approach with two well-referenced literature reviews (Benitti, 2012; Toh et al., 2016) to search papers in the ISI Web of Science. 64 references from backward snowballing and 97 citations from forward snowballing (a total of 161 papers) were found. With the above four inclusion criteria, five new papers were selected.

In the second round, seven papers (Avsec, Rihtarsic, & Kocijancic, 2014; Bers et al., 2014; Chen et al., 2017; Di Lieto et al., 2017; Kucuk & Sisman, 2017; Lindh & Holgersson, 2007; Rusk, Resnick, Berg, & Pezalla-Granlund, 2008) were identified as new seeds of the snowballing approach from the results of aforementioned searches. 125 references from backward snowballing and 120 citations from forward snowballing (a total of 245 papers) were examined. As a result, three new papers were singled out.

As the number of papers found in the second round was still large, we launched the third-round snowballing approach. This time we chose two papers (Chambers, Carbonaro, & Murray, 2008; Sullivan, 2008) as new seeds, and retrieved 120 papers (63 references from backward snowballing and 57 citations from forward snowballing). Only two new papers were selected for review.

After three rounds of snowballing approach, we found it difficult to find more qualified papers, and the repeated papers were increasing enormously (see Table 1). It indicated that the iterative process of snowballing approach could be ended.

Finally, as depicted in Table 1, 22 papers were selected as samples for the subsequent literature review.

Table 1
Summary of keyword search and snowballing approach.

Selection strategy	Papers resulting from the search	Selected
Keyword search	344	12(+3*)
First-round snowballing approach	161	5(+2*)
Second-round snowballing approach	245	3(+8*)
Third-round snowballing approach	120	2(+5*)
TOTAL	**870**	**22**

Note. * The number of papers selected before, which are not added to the total.

3. Results

3.1. How have robotics been incorporated into K-12?

In this section, we analyzed the 22 papers from the perspectives of sample groups, duration, robot types and robotics content knowledge involved. The detailed scheme is presented in Table A.1.

Among the 22 papers, 21 papers described the age or level of participants. Generally, robotics has been applied to education with a wide profile of students aged between 3 and 18, from kindergarteners to secondary school students. The frequency of sample groups selected for the 21 papers is shown in Fig. 1. The largest sample group is elementary school students (12 or 57.14%), followed by secondary school students (five or 23.81%) and kindergarteners (four or 19.05%).

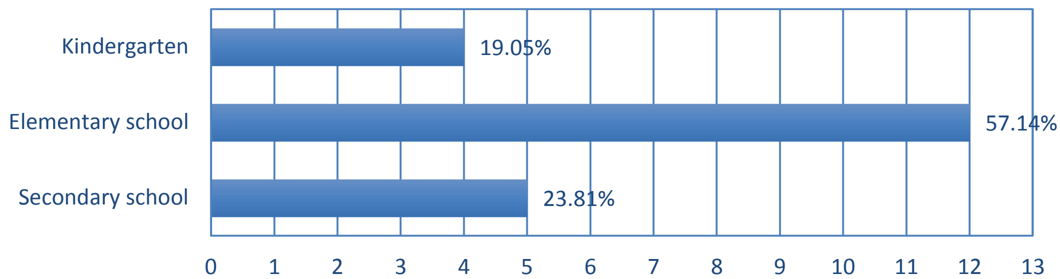


Fig. 1. Sample groups of these empirical studies

Among the 22 papers, 21 papers mentioned the number of participants. 10 or 47.62% of the papers recruited less than 40 participants (see Fig. 2). Only seven or 33.33% of the papers involved more than 80 participants. This indicates that the sample sizes were not large in the educational research of robotics. Promisingly, three papers distanced themselves from a small sample size: Lindh and Holgersson (2007) surveyed 696 students in a one-year regular robotic toys training; Menekse, Higashi, Schunn, and Baehr (2017) collected data from 366 adolescences that participated in a robotics tournament; and Peleg and Baram-Tsabari (2017) analyzed data from 391 questionnaires and interviews with 47 children and 20 parents.

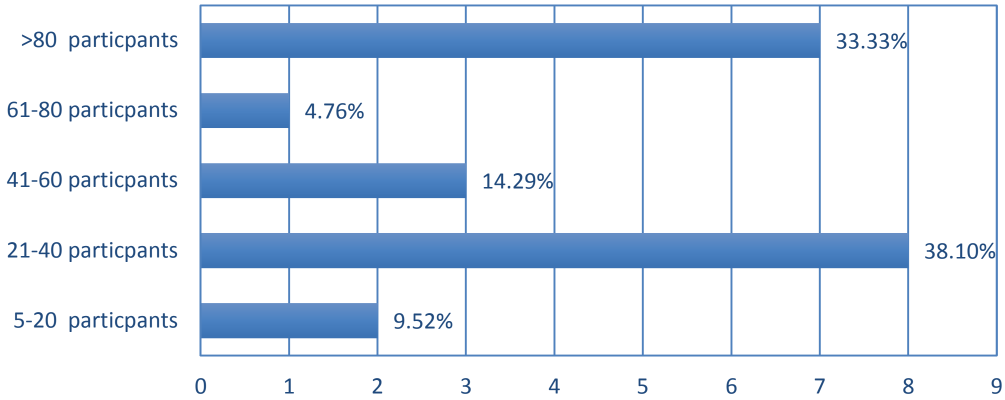


Fig. 2. Sample size of these empirical studies

Among the 22 papers, 16 papers discussed the duration of RE experiments. Three or 18.75% of the papers (Bers et al., 2014; Kucuk & Sisman, 2017; Master et al., 2017) controlled the time for less than one day. Seven or 43.75% of the papers (Avsec et al., 2014; Chambers et al., 2008; Di Lieto et al., 2017; Kazakoff & Bers, 2014; Mills, Chandra, & Park, 2013; Okita, 2014; Sullivan, 2008) conducted experiments between one and eight weeks. Two or 12.75% of the papers (Chen et al., 2017; Gomoll et al., 2016) set the research duration as 2-6 months. Only four or 25.00% of the papers (Jordan & McDaniel, 2014; Leonard et al., 2016; Lindh & Holgersson, 2007; Nemiro, Larriva, & Jawaharlal, 2017) conducted studies for more than one year. According to Fig. 3, 10 or 62.50% of the papers were endured less than two months. In other words, the duration of existing empirical studies is relatively short.

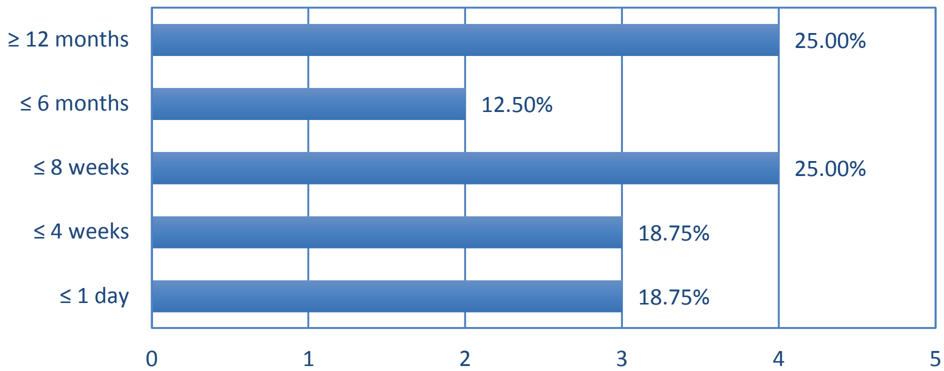


Fig. 3. Duration of these empirical studies

Considering the robot types used in these studies, the dominant position of LEGO is obvious (14 or 66.67%) (see Fig. 4). One paper combined LEGO and specially designed PicoCricket together in their robotics workshops (Rusk et al., 2008). In addition, one paper designed a customized robot for research purpose, i.e., “Pet” robot (Master et al., 2017). One did not mention the robot types (Peleg & Baram-Tsabari, 2017). Six papers used FicsherTechnik (Avsec et al., 2014), Nao (Chen et al., 2017), Bee-Bot (Di Lieto et al., 2017), iRobot Creates (Gomoll et al., 2016), Tiger Electronics “I-Cybie” (Jipson, Gülgöz, & Gelman, 2016) and Robotis Dream Level 1 (Kucuk & Sisman, 2017), respectively.

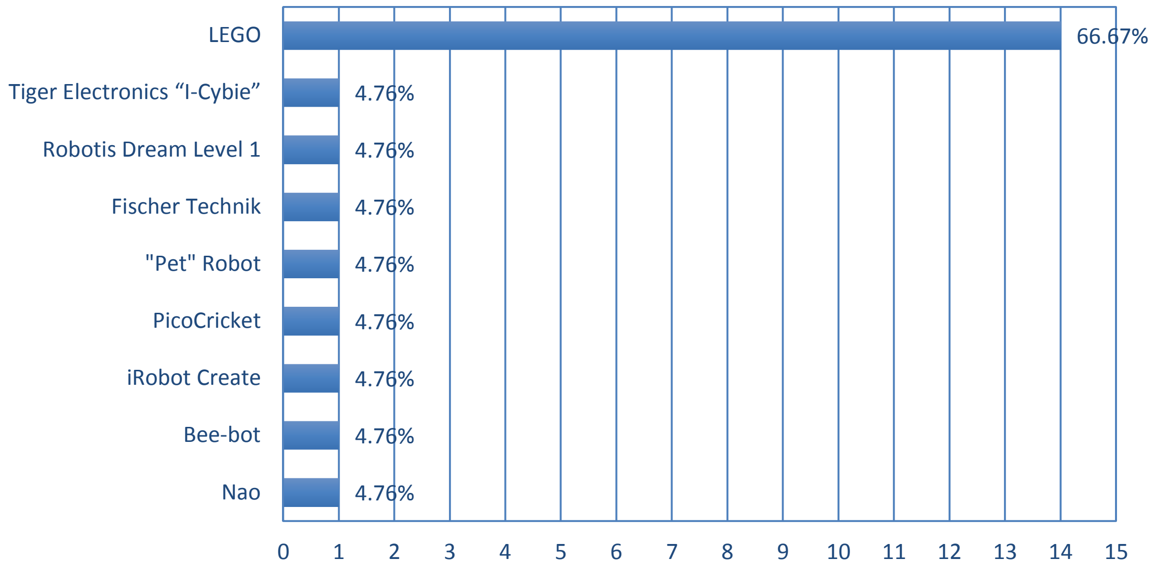


Fig. 4. Robot types in these empirical studies

The specific robotics content knowledge involved in the 22 papers is depicted in column 8 "*Content knowledge*" of Table A.1, which can be subdivided into seven categories: (1) structure and construction: gear, belt drive, etc.; (2) sensing: sound sensors, light sensors, etc.; (3) control: programming; (4) communication: wires, Wi-Fi; Bluetooth, etc.; (5) power and dynamical system: power, motor, Servo etc.; (6) concept and characteristics of robot; and (7) engineering design process. We delineate the frequency of robotics content knowledge involved in the 22 papers in Fig. 5. In general, most papers covered more than one category. 19 or 86.36% of the papers touched on the knowledge of control. 17 or 77.27% of the papers dealt with the structure and construction of robot. 13 or 59.09% of the papers considered the utilization of sensors. 10 or 45.45% of the papers conducted studies to learn the basic knowledge of power and dynamical system. Seven or 31.82% of the papers aimed to develop students' understanding of concept and characteristics of robot. Six or 27.27% of the papers covered the knowledge of communication. Three or 13.64% of the papers took engineering design process into consideration.

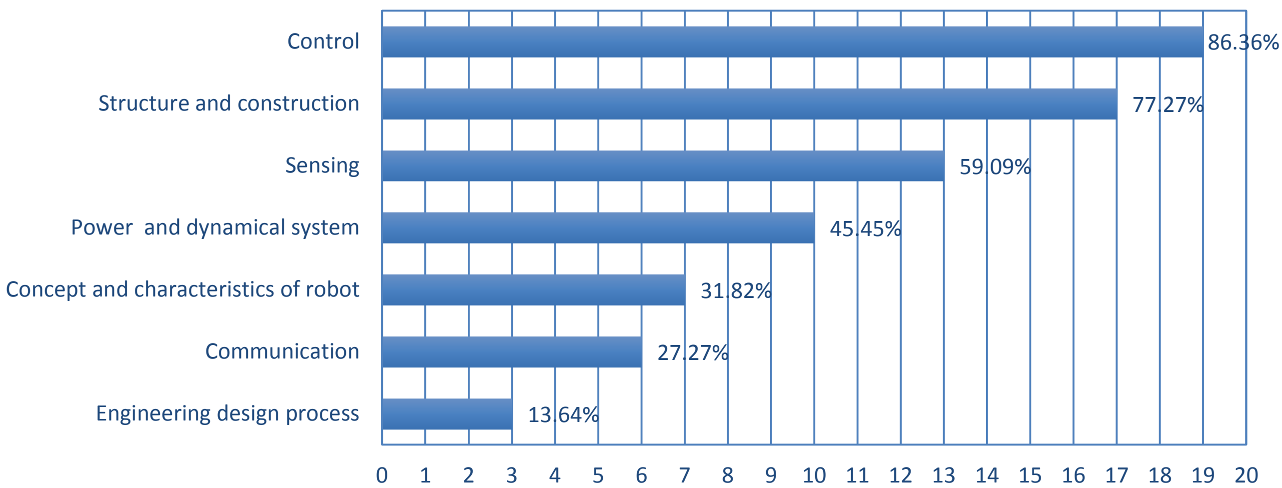


Fig. 5. Robotics content knowledge involved in these empirical studies

3.2. What intervention approaches are effective in teaching and learning robotics content knowledge?

In this section, we abstracted the study type, intervention approaches, major findings and measurement instruments reported in the 22 papers. The detailed scheme is presented in Table A.2.

As shown in Table A.2, we classified the 22 papers into non-experimental, quasi-experimental, and (true) experimental design according to the taxonomy of Benitti (2012) and Trochim and Donnelly (2006). For an experiment to be classified as a true experimental design, it must use random assignment to groups (random assignment is not the same thing as random selection of a sample from a population). If random assignment is not used, we have to ask a second question: does the design use either multiple groups or multiple waves of measurement? If the answer is yes, we would label it a quasi-experimental design. If not, we would call it a non-experimental design (Trochim & Donnelly, 2006). Among the 22 paper, the most popular research design is non-experiment (13 or 59.09%), followed by quasi-experiment (five or 22.73%) and experiment (four or 18.18%) (see Fig. 6). Among the nine experimental or quasi-experimental studies, five conducted with a control group.

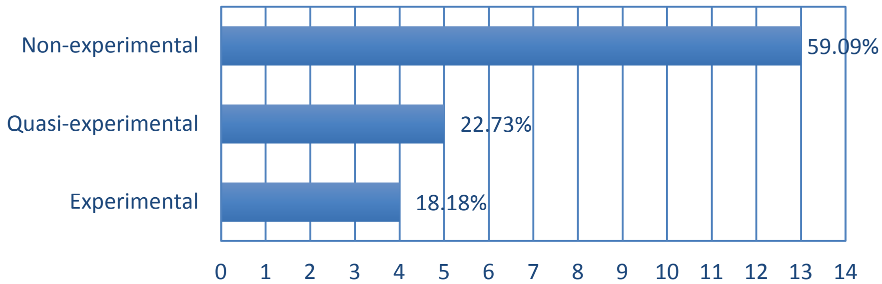


Fig. 6. Study type of these empirical studies

After analyzing the intervention approaches adopted in the 22 papers, three categories emerged: (1) pedagogical intervention, (2) robot allocation strategy and (3) programming environment (see Fig. 7). Among the 22 papers, 14 or 63.64% of the papers conducted RE with pedagogical interventions (e.g., problem-based learning, collaborative learning). Six or 27.27% of the papers adopted different robot allocation strategies in their RE treatments (e.g., LEGO versus non-LEGO). Two or 9.09% of the papers assigned students to different programming environments (e.g., text-based versus visual). Please refer to the columns "*Intervention*" and "*Study type*" in Table A.2 for more details about the intervention approaches adopted in each paper.

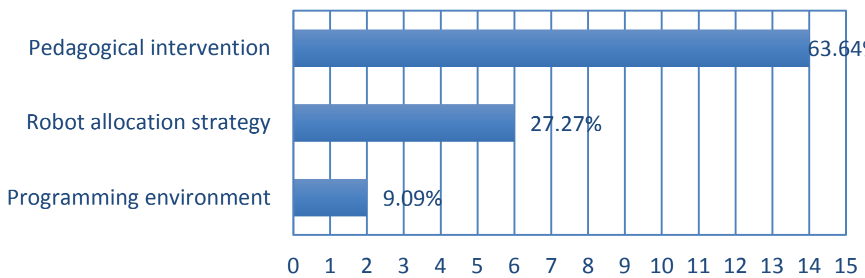


Fig. 7. Intervention approaches of these empirical studies

Generally, the results of the 22 papers showed students' multiple learning gains through RE (see the column "*Major findings*" in Table A.2). RE can help students benefit in: (1) understanding of concepts (e.g., programming, sensors, etc.); (2) change of attitudes (e.g., STEM motivation, self-efficacy, etc.); and (3) skills development (e.g., problem-solving, computational thinking, etc.). Nevertheless, this is not always the case, as there are indeed situations in which RE did not bring significant improvement in student learning. Chambers et al. (2008) reported that, though robotic sessions helped develop students' understanding of gear motion and function, the majority of students still failed to provide an accurate explanation about the concept of mechanical advantage. Leonard et al. (2016) found that students' attitudes toward STEM did not change significantly as a result of the study.

Gender difference is also an important issue related to RE. Among the 22 papers, Gomoll et al. (2016) and Master et al. (2017) found that gender discrepancy in STEM engagement and self-efficacy was large and persistent, thus RE treatments were conducted to bring girls into STEM in these two studies. Master et al. (2017) further reported that girls given programming experience reported higher technology interest and self-efficacy compared with girls without this experience and did not exhibit a significant gender gap relative to those of boys' interest and self-efficacy. Besides, Lindh and Holgersson (2007) revealed that there was no significant difference between boys and girls concerning the ability to build, program and more generally handle the LEGO material, however, gender difference was found in strategies of learning the material: boys more often were less willing to follow the instructions, whereas girls were more concentrated and intent on following the written task. Kucuk and Sisman (2017) indicated that male students needed guidance more often and they expressed feelings less often than female students in RE. Peleg and Baram-Tsabari (2017) found that girls expressed more negative attitudes towards robots than boys.

Seven measurement instruments were used in the 22 papers to evaluate students' performance: (1) observation, (2) questionnaire, (3) evaluation of artifacts, (4) verbal interview, (5) test/examination, (6) neuropsychological test battery and (7) self-report (see Table A.2). Most papers utilized more than one method for evaluation. As shown in Fig. 8, 13 or 59.09% of the papers used observation to obtain first-hand information of student performance. 11 or 50.00% of the papers conducted a questionnaire survey to investigate the learning outcomes. 10 or 45.45% of the papers implemented verbal interviews. Nine or 40.91% of the papers evaluated students' artifacts (e.g., robots, programs). Four or 18.18% of the papers used tests or examinations for evaluation. In addition to the above five conventional measurement instruments, two distinct instruments were found in two papers: Di Lieto et al. (2017) tested the learning effects of RE intervention with a neuropsychological test battery; Avsec et al. (2014) used a self-report instrument to measure students' expectations, satisfaction and outcomes of the content being learned/trained.

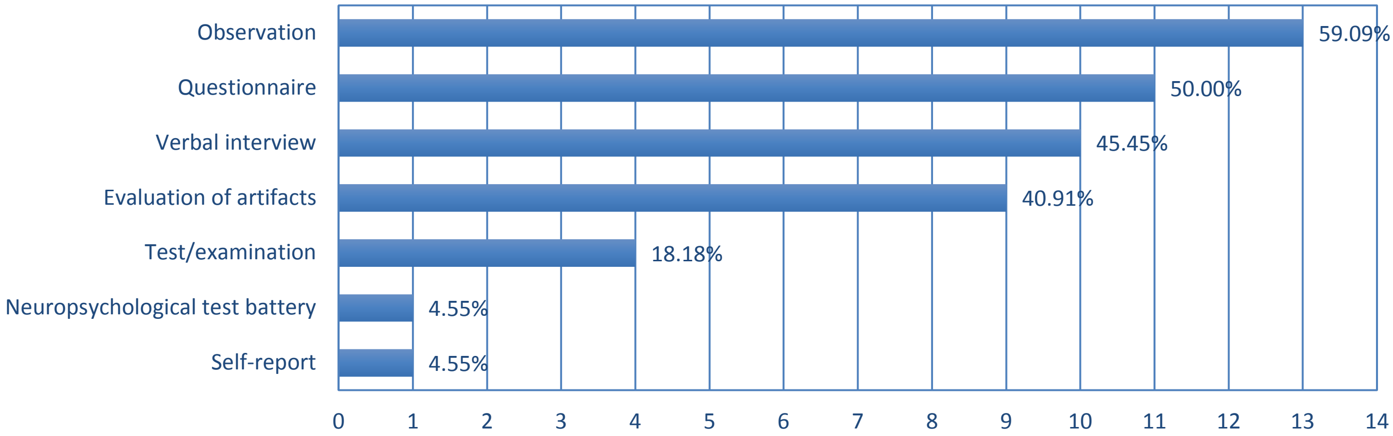


Fig. 8. Measurement instruments of these empirical studies

3.3. What implications for teaching are indicated by these empirical studies?

In order to provide both theoretical and practical reference for researchers and teachers alike, we extracted some useful instructional suggestions from the 22 papers.

3.3.1. Environment of RE activities

- *Large and open classroom*

Lindh and Holgersson (2007) suggested that a large space should be provided for the pupils to work. Nemiro et al. (2017) claimed that open physical layout is more beneficial to develop creative behaviors in students through robotics.

- *Appropriate hardware and software*

Bers et al. (2014), Kazakoff and Bers (2014), Rusk et al. (2008), Shih et al. (2013) and Sullivan (2008) advocated the importance of developmentally appropriate hardware and software specifically designed for young children. Okita (2014) found that initial knowledge building in less transparent programming environments can better prepare students to apply their knowledge to unfamiliar programming environments that require the students to solve problems using new principles.

- *Collaborative and creative learning atmosphere*

Hwang and Wu (2014) claimed that teachers should create collaborative learning environments and be prepared to work with different resources and tasks. Nemiro et al. (2017) suggested that it is important to guarantee a creative climate with a high level of classroom energy for developing creative behaviors in students through robotics.

3.3.2. Design of RE activities

- *Content design*

Avsec et al. (2014) suggested that professional content design and organization should be taken seriously to contribute to student satisfaction and course quality. Chen et al. (2017) and Lindh and Holgersson (2007) believed that the link between robotic activities and everyday life is important. Jordan and McDaniel (2014) advocated that robotics tasks and relational contexts should entail significant levels and sources of uncertainty. Kucuk and Sisman (2017) accentuated that high difficulty level robotics activities should be designed for students to develop their problem-solving abilities. Rusk et al. (2008) stated that robotics workshops should focus on themes, not just challenges and should try to incorporate art and engineering into projects.

- *Time allocation*

Bers et al. (2014) and Kucuk and Sisman (2017) demonstrated that more time should be devoted for children to build and explore the complex material. Chambers et al. (2008) suggested that children could be allotted time to examine the differences in robot construction and consequent behavior.

- *Activity organization*

Avsec et al. (2014) suggested that more collaborative activities should be organized to enhance learner-learner interaction. Gomoll et al. (2016), Kucuk and Sisman (2017), Master et al. (2017) and Rusk et al. (2008) listed out several strategies to engage more students into robotics: positive STEM experiences, multiple learning paths and deep engagement. Lindh and Holgersson (2007) indicated that the working groups should not be too big (maximum 2-3 pupils/LEGO box). Jordan and McDaniel (2014) argued that creating formal time to reflect on an introduction of and responses to uncertainty may enhance students' awareness of their behaviors related to managing uncertainty during collaborative problem solving. Kucuk and Sisman (2017) emphasized that teachers should motivate students to listen carefully when explaining robot bricks and the instruction process and should take gender difference into consideration. Meneske et al. (2017) suggested that it is important to ensure the collaboration quality among team members. Mills et al. (2013) argued that language and the use of tools play a dynamic and interactive role in the process of collaborative problem solving. Rusk et al. (2008) encouraged the organization of storytelling and exhibitions, rather than competitions in robotics workshops.

3.3.3. Pedagogy of RE activities

Chambers et al. (2008) demonstrated that a guided inquiry instructional approach is important during the early stages of developing a conceptual understanding of mechanical advantage. Sullivan (2008) highlighted that open-ended and extended inquiry can result in the use of thinking skills and science process skills in students. Gomoll et al. (2016) advocated the method of human-centered robotics to get girls interested and engaged in STEM. Leonard et al. (2016) encouraged teachers to use culturally relevant pedagogy and place-based education (i.e., students' culture and place) as springboards to maintain student interest in STEM.

3.3.4. Support of RE activities

- *Preparation*

Avsec et al. (2014) suggested that a technology training before robotics course may help increase students' confidence in performing tasks and in turn enhance student satisfaction and achievements.

- *Guidance*

Teachers must be able to support students and provide guidance during the whole robotics course (Kucuk & Sisman, 2017; Lindh & Holgersson, 2007). Students should be encouraged to share their ideas and experiences by asking questions to support their socialization (Kucuk & Sisman, 2017). In collaborative activities, because collaboration inevitably entails conflicts, teachers should guide students not only in preventing these conflicts but also in learning how to cope with conflicts and communicate and coordinate with others (Hwang & Wu, 2014). When designing problem-solving learning experiences, teachers should consider how to help students manage peer interaction, recognize the importance of interdependencies in their circumstances, and thereby effectively manage uncertainty with the help of their peers (Jordan & McDaniel, 2014). To develop creative behaviors in students, they should be encouraged to use both standard creativity techniques and techniques that are of personal value to them (Nemiro et al., 2017).

- *Feedback*

In robotics course, teachers should pay attention to students and provide feedback to them in a timely fashion (Avsec et al., 2014; Kucuk & Sisman, 2017; Sullivan, 2008).

4. Discussion

In this section, we generalize the main viewpoints of the 22 papers based on Section 3 and make a comparison with existing studies. Further, we point out future research perspectives of RE.

4.1. Incorporation of robotics into K-12

As shown in Fig. 9, despite the fact that we did not impose any restrictions on the publication time of reviewed papers, among the 22 SSCI journal papers, only four were published before 2013. Since then, the number of high-quality empirical studies has greatly increased and reached a peak of six papers in both 2014 and 2017. The trend line (see the red line in Fig. 9) shows the potential for future research and the rapid development of evidence-based research on teaching and learning robotics content knowledge in K-12. By contrast, in the study of Benitti (2012), under the inclusion criteria of peer-reviewed journal or conference papers, only 10 qualified papers (4 of these 10 papers were published in SSCI journals) were found, which were all published between 2006 and 2009. Therefore, we conclude that the growing trend of high-quality empirical studies on the teaching and learning of robotics content knowledge is significant in the latest five years.

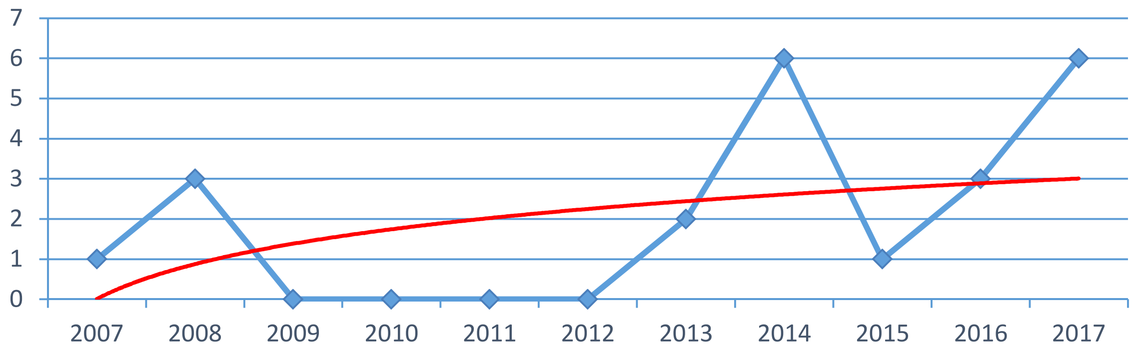


Fig. 9. Publication time of these empirical studies

As indicated in Section 3.1, among the 22 papers, many of them applied robotics with a short period and/or a small sample size. However, compared with the findings of Benitti (2012), some welcome changes emerged from this review: the number of papers with larger sample groups (more than 100 participants) and longer durations (more than two months) increased a lot. We propose that more RE experiments with a longer duration, preferably a longitudinal study, could be conducted to help ascertain the retention of learning effects as well as the transfer of knowledge in new learning environments. In the future, a larger sample size may make some of the effects that were non-significant in current studies become significant with greater power (Master et al., 2017).

Among the 22 papers, robots were used with a wide profile of K-12 students aged between 3 and 18. In contrast, Benitti (2012) reported that the participants in the 10 reviewed papers aged between 6 and 16, from 5th graders to 10th graders. This indicated that existing studies were conducted with more diverse sample groups in order to engage the younger (PreK-K) and older (11th–12th grade) students into robotics.

In spite of the variety of robotics kits available in the market now, most of the 22 papers (68.18%) used LEGO. Because of the limited school budget, it is impossible to provide each student this expensive robot equipment. Moreover, LEGO features as non-open robot kits, based on microcontrollers, and with a proprietary programing language (López-Rodríguez & Cuesta, 2016). Considering the above limitations of LEGO, it is more desirable to teach robotics content knowledge using open source hardware in the future. Arduino stands out from a large number of open source hardware and have been successfully used to develop students' skills, such as hardware configuration, OS administration, network configuration or high-level programming (López-Rodríguez & Cuesta, 2016; Martín-Ramos et al., 2017). More empirical studies are needed to demonstrate the practical value of these open-source robot kits in K-12, especially compared with commonly used LEGO.

4.2. Effective intervention approaches in teaching and learning robotics content knowledge

With respect to the study type, there is a dispute about its taxonomy in academia. For example, Toh et al. (2016) classified the study type taken by researchers in validating their use of robots into non-experimental and quasi-experimental, but it did not explain the rules of this classification in details. In this systematic review, we labeled the study type of the 22 papers into three categories according to Benitti (2012) and Trochim and Donnelly (2006), and presented statistical analysis in Section 3.2. Among the 22 papers, Avsec et al. (2014), Mills et al. (2013), and Sullivan (2008) claimed that it utilized an experimental research design, however, both random assignment and multiple groups or multiple waves of measurement were not found in them, thus the three studies were classified as non-experimental within our classification. Besides, two quasi-experimental study (Hwang & Wu, 2014; Shih et al., 2013) claimed to be a true experimental design, however, they did not use random assignment to groups. Table 2 presents a comparison between the study type claimed by these studies and the study type identified in this review. On the whole, the statistical results in Section 3.2 showed that there is a lack of quasi-experimental or experimental research design of RE in K-12, which is in line with Benitti (2012) and Toh et al. (2016). True experiment is generally the strongest of the three designs when establishing a cause-effect relationship. A non-experiment is generally the weakest in this respect. However, it does not mean that a non-experiment is the weakest of the three designs overall, but only with respect to internal validity or causal assessment (Trochim & Donnelly, 2006). Likewise, a control group is not always necessary in RE experiments, as we can gain lots of valuable information by a well-conducted non-experiment. Nonetheless, we may have a hard time establishing which of the things we observed are due to the robotics training rather than to other factors like students' prior knowledge or teacher's teaching style. In the future, quasi-experimental or experimental research design with a control group may better explore the cause-effect relationship that might exist between a certain RE treatment and its outcomes.

Table 2
The comparison of claimed and actual study type.

Paper	Claimed study type	Actual study type
Avsec et al. (2014)	Experiment	Non-experiment
Hwang & Wu (2014)	Experiment	Quasi-experiment
Mills et al. (2013)	Experiment	Non-experiment
Shih et al. (2013)	Experiment	Quasi-experiment
Sullivan (2008)	Experiment	Non-experiment

In terms of the intervention approaches used in the 22 papers, as shown in Section 3.2, the number of pedagogical interventions far surpassed that of the other two approaches. Nevertheless, the existing pedagogy adopted in RE (e.g., discovery learning, problem-based learning) failed in originality and was not targeted to the unique features of regular robotics curricula, i.e., multidiscipline, hands-on experience and integration (Qu & Wu, 2006). Based on the pedagogical technique of collaborative learning and pair programming (**Anonymous authors**, 2016; **Anonymous authors**, 2017), pair learning is an emerging and potential method in RE, which means that students collaborate in pairs to construct, build and program a robot under scripted but switchable roles. Future research can take advantages of this method and explore an effective design of pair learning in RE.

The overall results reported in the 22 studies indicated optimistic learning outcomes of students, including the improvements of knowledge, attitudes and skills. However, there were still some studies which reported inconclusive results of the role robot plays in education. Thus, future work could be conducted to provide more empirical evidence on the impacts of RE in K-12.

Although having made significant inroads into many traditionally male-dominated fields (e.g., biology, chemistry), female remained underrepresented in fields related to STEM (Cheryan, Master, & Meltzoff, 2015; Master, Cheryan, & Meltzoff, 2016), and this gender disparities can be traced back to primary school (Ceci & Williams, 2010). Among the 22 papers, gender differences were found in strategies of learning material (Lindh & Holgersson, 2007), behavioral patterns of teachers and students (Kucuk & Sisman, 2017) and attitudes toward robots (Peleg & Baram-Tsabari, 2017). However, Master et al. (2017) concluded that girls given programming experience did not exhibit a significant gender gap relative to boys' interest and self-efficacy. One possible explanation for this result may be that the duration of this experiment was relatively short (20 minutes) and the task assigned to students was relatively simple (using drag-and-drop visual programming to control the robot to move). In addition, Lindh and Holgersson (2007) also reported that they did not see any significant difference between boys and girls concerning the ability to build, program and more generally handle the LEGO material. It provided an evidence that like boys, girls have the same potential in RE, but they do not follow the usual male clichés of speed, power, competition, and destruction (Johnson, 2003). Due to the contradictory results on gender differences, future research could be conducted to further explore this important issue by rigorous empirical design.

It is apparent from this review that there is a need for a formal evaluation of RE. Self-designed measures, including observation, questionnaire, evaluation of artifacts, etc. are commonly-used measurement instruments in existing studies. However, these methods are summative in general, and lack of value for the development of education as well as students. One interesting finding is the utilization of self-report and neuropsychological test battery in two papers (Avsec et al., 2014; Di Lieto et al., 2017). These two evaluation methods provide the new path to conduct "true" process evaluation. **Anonymous authors** (2016) addressed that student's self-report was a more effective and efficient approach than interviews in-person or videotaped observations, because it can provide a helpful function for learning diagnosis and guidance in a low-cost. As to psychophysiological measurement, it is considered as an educational evaluation method with great potential in evidence-based research. For example, in studies on brain-computer interfaces, portable EEG is a type of psychophysiological measurement used to examine the relationships between mental and bodily processes, which can automatically measure the learner's attention and meditation levels in real-time in naturalistic classrooms settings (**Anonymous authors**, 2018). Since these two approaches are still in the early stages of RE research, there is also a need to conduct comprehensive experiments that adopted them in the future.

4.3. Instructional implications for K-12

The ultimate goal of educational research is to improve teaching and learning, yet the previous literature reviews (e.g., Benitti, 2012; Mubin et al., 2013; Toh et al., 2016) rarely focused on the instructional implications. From this perspective, we made a comprehensive analysis of the 22 papers included in this review and presented the instructional implications in details in Section 3.3.

On the practical level, teachers are the keystone to the implementation of RE. Berry, Remy, and Rogers (2016) found that many teachers could appreciate the benefits of robotics but they were uncomfortable with the technology aspect. Bers et al. (2014) reported that every participating teacher had about three trained assistants in the classroom to help troubleshoot technology issues, assess the children's progress, and provide one-on-one help as needed. In the future, professional teacher-training programs are needed to know more about what supports teachers need. For instance, the summer robotics workshop is a good program to provide opportunities for teachers to work with other like-minded individuals for integrating robotics into standard K-12 curriculum. Perhaps future work should attempt to assess the feasibility of implementing RE for one classroom teacher with minimal involvement of research assistants.

As an engineering activity, robotics is interdisciplinary and easily lends itself to collaboration

(Yuen et al., 2014). Among the 22 papers, a body of studies (e.g., Avsec et al., 2014; Hwang & Wu, 2014; Jordan & McDaniel, 2014; Lindh & Holgersson, 2007; Meneske et al., 2017) addressed students' collaborative learning of robotics in K-12. Currently, the ability to collaborate has been held with high virtue and is considered one of the 21st century skills (Trilling & Fadel, 2009). However, considering the commonly encountered problems of collaborative learning, such as free-rider issue (Puurtinen & Mappes, 2009), diffusion of responsibility (Karau & Williams, 1994), and negative groupthink effects (Kuhlen et al., 2005), there are also studies showing that competition is more effective to enhance students' motivation and learning (e.g., Lawrence, 2004). Future work could be done to incorporate the benefits of both collaboration and competition into RE, while minimizing their limitations. Therefore, coopetition, the hybrid behavior comprising simultaneous collaboration and competition, may be a good choice to mediate the debate of collaboration versus competition.

5. Conclusion

This study presents a systematic review of high-quality empirical studies on teaching and learning robotics content knowledge in K-12. The purpose of this study is twofold: (1) to summarize relevant empirical evidence on RE in K-12, and (2) to explore future research perspectives of RE. After a systematic search in online bibliographic database by keyword search and snowballing approach, 22 relevant papers were reviewed.

The participants of the 22 papers aged between 3 and 18, and the largest sample group was elementary school students. Nevertheless, more than half of the empirical studies adopted less than 80 participants and endured less than two months. Considering the robot types used in these papers, the dominant position of LEGO is obvious (14 or 66.67%). The specific robotics content knowledge mentioned in the 22 papers covered a wide range of the aforementioned seven categories. Future research could be conducted to explore the practical value of some open-source robotic kits (e.g., Arduino) to teach robotics content knowledge with a longer duration and a larger sample size.

We classified the 22 papers into non-experimental, quasi-experimental, and (true) experimental study according to the taxonomy of Benitti (2012) and Trochim and Donnelly (2006). As a result, 13 papers implemented a non-experimental research design. Among the nine experimental or quasi-experimental studies, only five conducted with a control group. In terms of the intervention approaches used in the 22 papers, the number of pedagogical interventions far surpassed that of other two approaches. However, the existing pedagogy adopted in RE failed in originality and was not targeted to the unique features of regular robotics curricula. Generally, the results of the 22 papers suggested multiple learning gains with RE, but there are indeed situations in which RE did not bring significant improvement in student learning. Thus, future work could be conducted to provide more empirical evidence with rigorous research design and targeted pedagogy to prove the impacts of RE in K-12. Seven measurement instruments were found in the 22 papers: (1) observation, (2) questionnaire, (3) evaluation of artifacts, (4) verbal interview, (5) test/examination, (6) neuropsychological test battery and (7) self-report. Generally, most studies utilized more than one method for evaluation. Apart from the first five conventional measures, the utilization of self-report and neuropsychological test battery provide new path to conduct "true" process evaluation. Since these two approaches are still in the early stages of RE research, there is, clearly, a need to conduct comprehensive experiment that adopted them in the future.

Because few literature reviews focused on the instructional implications, this study presented the instructional suggestions proposed in the 22 papers in detail in Section 3.3. We further divided these instructional suggestions into four themes: open environment, targeted design, appropriate pedagogy and timely support. On the practical level, teachers are the keystone to the implementation of RE. In the future, it will be necessary to know more about what supports teachers need in robotics course. Moreover, as collaboration and competition are both commonly-used way in RE, future work could be done to incorporate the benefits of both collaboration and competition into RE, while minimizing their limitations.

Overall, this study describes a promising picture of RE in K-12. It is hoped that this research can increase the knowledge on a yet largely unexplored, but significant topic, which is valuable both from a practical- and a research-oriented perspective.

Appendix

Table A.1
General information of the papers.

Paper	Participants	Age/Level	Duration	Robot	Content knowledge
Avsec et al. (2014)	105	13-18	One week in 2012 One week in 2013	FicsherTechnik	Fundamental structure of simple sensors. The influence of direction of the electric current through the DC motor. Gear units and other mechanisms. Algorithms with loop-back operation.
Bers et al. (2014)	53	4.9-6.5	Approximately 20 hours of classroom work	LEGO	Engineering Design Process. Knowledge of robotics (basic ideas, wires, motors, and construction). Choosing, sequencing, looping, and branching programs. Sensors.
Chambers et al. (2008)	22	9-10	Six sessions over six weeks	LEGO Mindstorms	Gear motion and function. Mechanical advantage.
Chen et al. (2017)	37	5th grade	Weekly session over six months	Nao	Writing programs around specific topics, such as basic actions, voice recognition, tactile sensors, walking motion, animation, and key computer science concepts.
Di Lieto et al. (2017)	12	5-6	Twice a week for six weeks	Bee-Bot	Familiarization of Bee-Bot use (programming characteristics and main buttons). The steps of programming. Programming Bee-Bot following different and contradictory rules on a more abstract map.
Gomoll et al. (2016)	22	6th-12th grade	1-2 hours each Friday for 11 weeks	iRobot Creates	Designing and building telepresence robots — remote-controlled robots equipped to navigate a space using motor, programs, wheels, sensor data, and a camera. Engineering Design Process.
Hwang & Wu (2014)	48	6th grade	N/A	LEGO	Controlling robots and sending messages to PDAs via wireless internet access and then to robots (via Bluetooth).
Jipson et al. (2016)	36	3-5	N/A	Tiger Electronics "I-Cybie"	Biological, psychological, sensory, and artifact properties, as well as their use of other subtle indicators of ontological status.
Jordan & McDaniel (2014)	24	5th grade	Every Friday morning across the school year	LEGO Mindstorms	Building robots by using basic programming techniques to enable locomotion and the activation of sensors and motors. Designing its own robot to address an environmental problem the group identified.
Kazakoff & Bers (2014)	34	4.5-6.5	No more than three weeks	LEGO Mindstorms	Building and programming LEGO robot. Understanding of robotics concepts.
Kucuk & Sisman (2017)	18	8-11	Two and three hours	Robotis Dream Level 1	Introduction of the parts of the robot education kit. Combining robotics parts. Use of switch, geared motor, liion battery and led parts. The principle of center of gravity
Leonard et al. (2016)	76	5th-8th grade	Three years	LEGO EV3	Putting basic pieces together to make a robot. Using simple code to move robots. Learning to use sensors, such as color and touch.
Lindh & Holgersson (2007)	696	12-16	Two hours per week for one year	LEGO Dacta	Constructing the robotics. Programming the computers. Loading programs to the robotics via an infrared beam.
Master et al. (2017)	96	6	20 minutes	Specially designed "Pet" robot	Using a smartphone to program the robot.
Menekse et al. (2017)	366	10-13	N/A	LEGO	Designing, building, and programming robots using the knowledge of sensors, power, and communication. Devising an innovative solution for a research project.
Mills et al. (2013)	24	8.5-9.5	An intensive	LEGO	Building and programming LEGO robots on wheels to move to a target destination. New

			seven-week block	Mindstorms	technical, mathematical, and science vocabulary and key concepts.
Nemiro et al. (2017)	194	4th-6th grade	Weekly observations over three years	LEGO	Robotics design, construction, and programming. Ultrasonic sensors, and motors.
Okita (2014)	41	9-11	One-hour session for 10 consecutive days	LEGO Mindstorms NXT 2.0	Programming robot movements using abstract concepts of speed, distance and direction. Light sensors which can scan handwritten program code and execute action.
Peleg & Baram-Tsabari (2017)	315	1th-4th grade	N/A	N/A	Each robot needs to have a motor, sensors and a computer. Robots can help humans in numerous ways. Children can design simple robots. Robots are not human—they need not look like people and they cannot replace human love.
Rusk et al. (2008)	N/A	N/A	N/A	LEGO; PicoCricket	Mechanical construction, programming; art and engineering, light sensors, touch sensors, sound sensors; motors, colored lights, music-making devices; bricks, gears, and axles et al.
Shih et al. (2013)	102	5th grade	N/A	LEGO NXT	Programming of LEGO robotics.
Sullivan (2008)	26	11-12	Two 3-week sessions	LEGO Mindstorms	Designing and building a structure. Programming the structure to accomplish a specific task using the knowledge of light sensors, motor, and tower.

N/A – Not available.

Table A.2
Study design and major findings.

Paper	Intervention	Study type [a]	Measurement instruments	Major findings	
				Proved results	**Non-proved results**
Avsec et al. (2014)	Non-experimental: Discovery learning; Problem-based learning.	X O	Self-report.	Students have significantly positive perceptions toward using open learning of robotics as a learning-assisted tool.	It is unclear that whether specific variations in teaching are particularly salient to the satisfaction of girls, and whether these results could be explained by gender differences.
Bers et al. (2014)	Non-experimental: Teacher-guided whole-class approach.	X O	Evaluation of artifacts; Verbal interview.	Kindergartners were both interested in and able to learn many aspects of robotics, programming, and computational thinking with the TangibleK curriculum.	There were no statistically significant differences between children's score on a concept in the introductory activity and the score on that same concept in the final project.
Chambers et al. (2008)	Non-experimental: Semi-structured and guided scientific inquiry approach.	O X O	Verbal interview.	The robot sessions helped develop the students' understanding of gear motion and function.	Most children were unable to accurately explain the concept of mechanical advantage.
Chen et al. (2017)	Experimental: EG1: text-based programming language; EG2: visual programming language.	R O X1 O R O X2 O	Test/examination.	Robotics curriculum did help improve students' CT, especially in the context of robotics programming.	It is unclear that to what extent did the curriculum factor into the achieved improvement on algorithmic thinking. Students overall did not gain as much in dealing with CT applied to everyday life as that to programming robotics.
Di Lieto et al. (2017)	Quasi-experimental: EG: Robotics activities with Bee-Bot; CG: daily school activities.	N O X O N O O	Neuropsychological test battery; Evaluation of artifacts; Questionnaire.	A significant improvement in both visuo-spatial working memory and inhibition skills after the robotics period, with a significant effect also on robot programming skills.	N/A
Gomoll et	Non-experimental:	O X O	Verbal interview;	Opportunities to personalize students' robots and	It is unclear that how to harness multiple forms of

al. (2016)	Human-centered robotics; Problem-based learning.		Observation; Evaluation of artifacts.	feedback from peers and facilitators were important motivators.	leadership and engagement without marginalizing students with different working preferences.
Hwang & Wu (2014)	Quasi-experimental: EG1: three students to three robots; EG2: three students to two robots; EG3: two students to three robots.	N X1 O N X2 O N X3 O	Questionnaire; Observation.	Three collaborative strategies for solving problems emerged from the three scenarios: independent-control, mutual-control and coordinator-directed. The mutual-control collaborative strategy increased the frequency of behavioral interactions.	It is unclear that whether the strategy of coordinator helped collaboration. The negative correlation that expected between coordination and robot conflict was not supported by the statistical analysis.
Jipson et al. (2016)	Non-experimental: Parent–child conversation.	X O	Verbal interview; Questionnaire; Observation.	Parent talk in the play session had the most influence on children's reasoning when the properties under consideration were less well-established in children's thinking and/or not easily identified by visual cues (i.e., psychological and sensory).	It is not entirely clear why the property projection task and parent-child activity session yielded distinct results.
Jordan & McDaniel (2014)	Non-experimental: Direct teacher instruction reinforced with online support; Collaborative learning.	X O	Observation; Verbal interview; Evaluation of artifacts.	Three socially supportive peer responses and two unsupportive peer responses were identified. Peer interaction was influential because students relied on supportive social response to enact most of their uncertainty management strategies.	When one is dealing with an internal state unobservable by direct means, any interpretation is tentative, even when supported by participants' interpretation of their own actions, as there are multiple ways to interpret any social interaction
Kazakoff & Bers (2014)	Non-experimental: One-to-one sessions with researchers.	O X O	Verbal interview; Observation; Evaluation of artifacts; Questionnaire.	Young children can improve sequencing skills through learning to program robots with developmentally appropriate tools.	It is unclear that whether these increases temporary or long-term, and whether the exposure to computer programming and robotics in early childhood would increase interest in STEM subjects or the likelihood of STEM career choices.
Kucuk & Sisman (2017)	Non-experimental: One-to-one robotics instruction.	X O	Observation.	Students' assembling bricks, sharing ideas and experiences, and teachers' providing guidance and asking questions were the most frequent behaviors. Gender differences did exist in behavioral patterns of students and teachers.	N/A
Leonard et al. (2016)	Quasi-experimental: EG1: robotics only; EG2: gaming only; EG3: robotics/gaming.	N O X1 O N O X2 O N O X3 O	Questionnaire; Observation; Evaluation of artifacts.	Students who participated in blended robotics/gaming clubs had significantly higher self-efficacy scores on the construct of videogaming. Students who participated in holistic game development had higher CT ratings.	Students' self-efficacy significantly declined on the construct of computer use. Student attitudes toward STEM did not change significantly.
Lindh & Holgersson (2007)	Experimental: EG: LEGO; CG: non-LEGO.	R O X O R O O	Observation; Verbal interview; Questionnaire; Test/examination.	There are significant positive effects of LEGO for sub groups of pupils.	There is no obvious over-all effect of LEGO. There is no significant difference between boys and girls concerning the ability to build, program and more generally handle the LEGO material.
Master et al. (2017)	Experimental: EG1: "robot" activity; EG2: "storytelling" activity;	R X1 O R X2 O R O	Questionnaire.	Girls given programming experience reported higher technology interest and self-efficacy compared with girls without this experience.	Girls given programming experience did not exhibit a significant gender gap relative to boys' interest and self-efficacy.

	CG: no activity.				
Menekse et al. (2017)	Non-experimental: Competition; Collaboration; Problem-based learning.	X O	Evaluation of artifacts; Verbal interview; Observation; Questionnaire.	Collaboration quality was a good predictor of robotics team performance. The cumulative amount of team experience was significantly related to collaboration quality.	It remains unclear exactly what specific features of FFL competition lead to the outcomes observed: Was it the design of the game board tasks, or perhaps the cultural emphasis on teamwork as a core value among participants?
Mills et al. (2013)	Non-experimental: Collaborative problem solving with robotics.	X O	Observation; Verbal interview; Evaluation of artifacts.	In the context of collaborative robotics problem solving in school, language and the use of tools play a dynamic and interactive role in the learning process.	N/A
Nemiro et al. (2017)	Non-experimental: problem-based robotics program; Authentic learning.	X O	Observation.	The SRI was shown to develop creative behavior in the students through the myriad of creative robot products generated.	N/A
Okita (2014)	Experimental: EG1: high-transparency environments learned visual programming; EG2: low-transparency environments learned syntactic programming.	R X1 O X2 O R X2 O X1 O	Test/examination.	The posttest revealed the benefits of initial learning in low-transparency environments as students performed better on repeated and new inferential problems across virtual and physical platforms	Students who experienced initial knowledge building in high-transparency environments showed less adaptability and lower performance on unfamiliar problems from low-transparency environments.
Peleg & Baram-Tsabari (2017)	Quasi-experimental: EG: watched the play "Robot and I"; CG: did not watch the play "Robot and I".	N X O N O	Questionnaire; Verbal interview.	Explicit but not implicit learning goals were decoded successfully. Gender differences did exist in certain questionnaire items regarding attitudes or interest.	Regarding the learning goals of "Children can be inventive and design things." no significant difference was found between EG and CG.
Rusk et al. (2008)	Non-experimental (anecdotal case studies): Theme-based learning; Project-based learning.	X O	Evaluation of artifacts; Observation.	Providing multiple paths can engage young people with diverse interests and learning styles into robotics.	N/A
Shih et al. (2013)	Quasi-experimental: EG: LEGO NXT; CG: information technology.	N O X O N O O	Test/examination; Questionnaire.	The learning effects of students instructed with LEGO NXT in the teaching are superior to students instructed with information technology integrated in the teaching.	There were no significant impacts of perceived ease of use on attitude and behavioral intention. Two hypotheses of perceived usefulness had no significant impacts on behavioral intention. The impact of behavioral intention on learning effects did not reach significance.
Sullivan (2008)	Non-experimental: Direct instruction; Open-ended, student-directed inquiry.	O X O	Observation; Questionnaire.	Students used seven out of eight thinking skills and science process skills described in the coding scheme. Students improved their systems understanding.	The only thinking skill that was not used by the majority of the students was computation.

[a]: N = nonequivalent group; O = measures / evidence; R = random assignment; X = treatment.
EG – Experimental group; CG – Control group.
N/A – Not available.

References

Avsec, S., Rihtarsic, D., & Kocijancic, S. (2014). A predictive study of learner attitudes toward open learning in a robotics class. Journal of science education and technology, 23(5), 692-704.

Benitti, F. B. V. (2012). Exploring the educational potential of robotics in schools: A systematic review. Computers & Education, 58(3), 978-988.

Berry, C. A., Remy, S. L., & Rogers, T. E. (2016). Robotics for all ages: a standard robotics curriculum for k-16. IEEE Robotics & Automation Magazine, 23(2), 40-46.

Bers, M. U., Flannery, L., Kazakoff, E. R., & Sullivan, A. (2014). Computational thinking and tinkering: Exploration of an early childhood robotics curriculum. Computers & Education, 72, 145-157.

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

Chambers, J. M., Carbonaro, M., & Murray, H. (2008). Developing conceptual understanding of mechanical advantage through the use of LEGO robotic technology. Australasian Journal of Educational Technology, 24(4), 387-401.

Chen, G., Shen, J., Barth-Cohen, L., Jiang, S., Huang, X., & Eltoukhy, M. (2017). Assessing elementary students' computational thinking in everyday reasoning and robotics programming. Computers & Education, 109, 162-175.

Cheryan, S., Master, A., & Meltzoff, A. N. (2015). Cultural stereotypes as gatekeepers: increasing girls' interest in computer science and engineering by diversifying stereotypes. Frontiers in psychology, 6, 1-8.

Di Lieto, M. C., Inguaggiato, E., Castro, E., Cecchi, F., Cioni, G., Dell'Omo, M., ... & Dario, P. (2017). Educational robotics intervention on executive functions in preschool children: a pilot study. Computers in Human Behavior, 71, 16-23.

Flynn, B. B., Sakakibara, S., Schroeder, R. G., Bates, K. A., & Flynn, E. J. (1990). Empirical research methods in operations management. Journal of operations management, 9(2), 250-284.

Gomoll, A., Hmelo-Silver, C. E., Šabanović, S., & Francisco, M. (2016). Dragons, Ladybugs, and Softballs: Girls' STEM Engagement with Human-Centered Robotics. Journal of Science Education and Technology, 25(6), 899-914.

Hwang, W. Y., & Wu, S. Y. (2014). A case study of collaboration with multi-robots and its effect on children's interaction. Interactive Learning Environments, 22(4), 429-443.

Jipson, J. L., Gülgöz, S., & Gelman, S. A. (2016). Parent–child conversations regarding the ontological status of a robotic dog. Cognitive Development, 39, 21-35.

Johnson, J. (2003). Children, robotics, and education. Artificial Life and Robotics, 7(1), 16-21.

Jordan, M. E., & McDaniel Jr, R. R. (2014). Managing uncertainty during collaborative problem solving in elementary school teams: The role of peer influence in robotics engineering activity. Journal of the Learning Sciences, 23(4), 490-536.

Karau, S. J., & Williams, K. D. (1994). Social loafing: A meta-analytic review and theoretical integration. Journal of Personality & Social Psychology, 65(4), 681-706.

Kazakoff, E. R., & Bers, M. U. (2014). Put your robot in, put your robot out: Sequencing through programming robots in early childhood. Journal of Educational Computing Research, 50(4), 553-573.

Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering–a systematic literature review. Information and software technology, 51(1), 7-15.

Kucuk, S., & Sisman, B. (2017). Behavioral patterns of elementary students and teachers in one-to-one robotics instruction. Computers & Education, 111, 31-43.

Kuhlen, R., Griesbaum, J., Jiang, T., König, J., Lenich, A., Meier, P., ... & Semar, W. (2005). K3–an e-learning forum with elaborated discourse functions for collaborative knowledge management. In the proceeding of E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education (pp. 2981-2988).

Lawrence, R. (2004). Teaching data structures using competitive games. IEEE Transactions on Education, 47(4), 459-466.

Leonard, J., Buss, A., Gamboa, R., Mitchell, M., Fashola, O. S., Hubert, T., & Almughyirah, S. (2016). Using Robotics and Game Design to Enhance Children's Self-Efficacy, STEM Attitudes, and Computational Thinking Skills. Journal of Science Education and Technology, 25(6), 860-876.

Lindh, J., & Holgersson, T. (2007). Does LEGO training stimulate pupils' ability to solve logical problems?. Computers & education, 49(4), 1097-1111.

López-Rodríguez, F. M., & Cuesta, F. (2016). Andruino-A1: low-cost educational mobile robot based on Android and Arduino. Journal of Intelligent & Robotic Systems, 81(1), 1-14.

Martín-Ramos, P., Lopes, M. J., da Silva, M. M. L., Gomes, P. E., da Silva, P. S. P., Domingues, J. P., & Silva, M. R. (2017). First exposure to Arduino through peer-coaching: Impact on students' attitudes towards programming. Computers in Human Behavior, 76, 51-58.

Master, A., Cheryan, S., & Meltzoff, A. N. (2016). Computing whether she belongs: Stereotypes undermine girls' interest and sense of belonging in computer science. Journal of Educational Psychology, 108(3), 424-437.

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

Menekse, M., Higashi, R., Schunn, C. D., & Baehr, E. (2017). The Role of Robotics Teams' Collaboration Quality on Team Performance in a Robotics Tournament. Journal of Engineering Education, 106(4), 564-584.

Mills, K. A., Chandra, V., & Park, J. Y. (2013). The architecture of children's use of language and tools when problem solving collaboratively with robotics. The Australian Educational Researcher, 40(3), 315-337.

Mubin, O., Stevens, C. J., Shahid, S., Al Mahmud, A., & Dong, J. J. (2013). A review of the applicability of robots in education. Technology for Education and Learning, 1, 1-7.

Nemiro, J., Larriva, C., & Jawaharlal, M. (2017). Developing Creative Behavior in Elementary School Students with Robotics. The Journal of Creative Behavior, 51(1), 70-90.

Okita, S. Y. (2014). The relative merits of transparency: Investigating situations that support the use of robotics in developing student learning adaptability across virtual and physical computing platforms. British Journal of Educational Technology, 45(5), 844-862.

Peleg, R., & Baram-Tsabari, A. (2017). Learning Robotics in a Science Museum Theatre Play: Investigation of Learning Outcomes, Contexts and Experiences. Journal of Science Education and Technology, 26(6), 561-581.

Puurtinen, M., & Mappes, T. (2009). Between-group competition and human cooperation. Proceedings of the Royal Society of London B: Biological Sciences, 276(1655), 355-360.

Qu, Z., & Wu, X. (2006). A new curriculum on planning and cooperative control of autonomous mobile robots. International Journal of Engineering Education, 22(4), 804-814.

Rusk, N., Resnick, M., Berg, R., & Pezalla-Granlund, M. (2008). New pathways into robotics: Strategies for broadening participation. Journal of Science Education and Technology, 17(1), 59-69.

Shih, B. Y., Chen, T. H., Wang, S. M., & Chen, C. Y. (2013). The exploration of applying LEGO NXT in the situated science and technology learning. Journal of Baltic Science Education, 12(1), 73-91.

Sullivan, F. R. (2008). Robotics and science literacy: Thinking skills, science process skills and systems understanding. Journal of Research in Science Teaching, 45(3), 373-394.

Toh, L. P. E., Causo, A., Tzuo, P. W., Chen, I., & Yeo, S. H. (2016). A review on the use of robots in education and young children. Journal of Educational Technology & Society, 19(2), 148-163.

Trilling, B., & Fadel, C. (2009). 21st century skills: Learning for life in our times. John Wiley & Sons.

Trochim, W. M. K., & Donnelly, J. P. (2006). Research methods knowledge base (3rd ed.). Mason, OH: Thomson. Retrieved form http://www.socialresearchmethods.net/kb/

Wohlin, C. (2014). Guidelines for snowballing in systematic literature studies and a replication in software engineering. In Proceedings of the 18th international conference on evaluation and assessment in software engineering (p. 38). ACM.

Yuen, T. T., Boecking, M., Tiger, E. P., Gomez, A., Guillen, A., Arreguin, A., & Stone, J. (2014). Group tasks, activities, dynamics, and interactions in collaborative robotics projects with elementary and middle school children. Journal of STEM Education: Innovations and Research, 15(1), 39-45.

Anonymous authors. (2018). This paper is published in Computers in Human Behavior.

Anonymous authors. (2017). This paper is published in Journal of Educational Technology & Society.

Anonymous authors. (2016). This paper is published in Computers in Human Behavior.

Anonymous authors. (2016). This paper is published in Journal of Educational Computing Research.

Highlights

- The growing trend of research on robotics education (RE) in K-12 is significant.
- RE has the great educational potential in K-12.
- RE calls for process evaluation methods and featured pedagogical interventions.
- More empirical studies with rigorous design should be done for the future of RE.

Funding

This work was supported by the “2011 Plan of Jiangsu: Collaborative Innovation Center for Strengthen Moral Education and Cultivate People, Nanjing Normal University”, and the “Priority Academic Program Development of Jiangsu Higher Education Institutions”.