Computer Science in Schools: A Literature Mapping of Professional Development for In-Service Teachers

Majid Rouhani
Department of Computer Science
Norwegian University of Science and Technology
Trondheim, Norway
majid.rouhani@ntnu.no

Monica Divitini
Department of Computer Science
Norwegian University of Science and Technology
Trondheim, Norway
divitini@ntnu.no

Amir Massoud Hashemi
Department of Electrical engineering and Mathematical sciences
Western Norway University of Applied Science
Bergen, Norway
amir.massoud.hashemi@hvl.no

2022 IEEE Global Engineering Education Conference (EDUCON) | 978-1-6654-4434-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/EDUCON52537.2022.9766599

***Abstract*—During the last few decades, there has been a growing need for Professional Development of in-service teachers in Computer Science. This paper presents a systematic mapping of the research in this area published between 2010 and 2020. The study's goal is to map existing literature, understanding how research is evolving and identifying gaps that can prompt new research. The literature mapping is based on the analysis of 206 articles collected from various online databases, then selected according to defined inclusion/exclusion criteria. The study investigates the type of intervention/study, school level, geographical location, connection to curricula, learning objectives, phase of the research, size of the study, and type of collaboration. From the literature mapping, Professional Development of in-service teachers of Computer Science emerges as a rapidly growing and dynamic research area. However, some threats are connected to fragmentation of the research and the need for more cooperation to increase inclusiveness and international collaboration.**

***Index Terms*—Professional Development, In-service teachers, Computer Science, Literature mapping, K-12**

I. Introduction

A lot of attention is being paid to recent and upcoming computer science curriculum revisions around the world. Two examples of audits and analyses of national-level curricula models are [1], and [2], focusing on the USA and Europe. They both identify access to qualified teachers as a challenge to the successful implementation of the new curricula. With the reforms now in place, there is a pressing need for teachers to get Professional Development (PD) in various formats [3].

Some examples of the type of intervention that have been adopted are workshops [4], summer programs [5], courses [6], [7], and in-school activities [8]. Online PD for Computer Science (CS) teachers, such as [9], [10], has become a popular approach to teachers PD and might be a potential way to give training to CS teachers.

Although there has been some research on teachers' Professional Development in the last decade, there is little overview of the studies and the overall global development of the field. We examined the research within PD aimed at in-service teachers to understand how it is evolving and identified some literature review papers but did not find any literature mapping showing the development over time. Reference [11] did a literature review on computational thinking (CT) from different perspectives in K-12 and came to an agreement on CT as a cultural method, as well as a shared approach to a refined working definition. Reference [12] conducted a study of CS learning and teaching resources for K-12 computer curricula. Their findings show a strong preference for high-quality resources. They did, however, identify several critical gaps. Reference [13] did a review on CS teacher PD in the US. However, this review was for studies published between 2004 and 2014.

The primary goal of a systematic mapping study is to present an overview of a study topic and determine the amount and variety of available research and findings [14]. Systematic reviews, on the other hand, are concerned with gathering and synthesizing evidence [15]. We based our systematic mapping approach on the work of [15], [16]. Our literature study aims to map research on PD for in-service teachers of CS in the period 2010-2020. This mapping provides a bird's-eye view of current research, helping the community to understand how research in this area is evolving and identifying gaps in recent research that can prompt new research.

The next section (II) presents the motivation and research questions defined for this study. Then, we present our methodology in Section III, and answer the research questions in Section IV. In Section V, we discuss the main findings in this study and present the conclusions and further work in Section VI.

II. Motivation, research questions and scoping

Preparing teachers in this area has become more critical as the focus on introducing primary and secondary pupils to CS has expanded. Several national curricula support this renewed interest, policy reforms and standards [17] such as CS Teachers Association Standards [18] and National Curriculum for England [19].

Teachers may participate in different types of interventions, such as workshops [20], courses [6], projects, or incorporate these types of activities within their regular duties. We see many kinds of interventions and want to know which form is most common. We ask the research question (RQ.1): What kinds of interventions are reported in the literature?

What we think is natural to investigate further is then which school level these interventions are aimed at. We can see that PDs are provided in all levels ranging from preschool [21], after-school setting [22] to precollegiate [23]. We want to know what schools levels are in focus and ask the research question (RQ.2): At what school level is research done?

We find many articles from specific areas of the world, while many other areas appear to be absent. Classification of studies based on the geographical area can help us see the gap in research on PD worldwide. Therefore, we defined the research question (RQ.3): In which geographical areas is the research done?

Another question that arises is to what extent the intervention is rooted in the national curricula. More specifically, we look for which subject(s) the intervention supports. For example, an intervention could be connected to STEM-related subjects [24] and/or English [25]. This will also help us understand where there is a gap and where there is a focus. Therefore, we defined the research question (RQ.4): How are interventions connected to the national curriculum?

We see many types of PD with different purposes ranging from integrating computational thinking into science lessons [26], assessment of PD [27], interdisciplinary training in programming [28], and much more. Classification of learning objectives may help us identify the gaps in PD in CS. We have defined the research question (RQ.5): Which are the learning objectives of the interventions?

The research that is carried out can be in different phases. It may be research that aims to inform the design of PD interventions and programs [29]. Results from such studies can be used later when designing and implementing an intervention. There may also be research done through the design of an intervention or about evaluating an intervention [30]. To understand the investigation's phase and which phases are more used, we defined this research question (RQ.6): At what phase is the research done?

Research studies can be conducted on smaller or larger groups. We find studies of very different sizes, ranging from small scale workshops [20] to large MOOC [31]. We want to know more about the size of studies and therefore asked the question: (RQ.7): What is the size of studies presented in the literature? This classification refers to the number of teachers that are involved.

We find studies where different actors are involved either within the same institution [32] or across different organizations [33]. It might be interesting to see if PD interventions mobilize different actors in the communities (and then which ones) and if this type of cooperation has increased or decreased. Therefore, we seek to answer this research question (RQ.8): What type of collaboration is connected to the research?

III. Method Description

In this study, we use systematic mappings (SM). Systematic mappings are not the same as systematic literature reviews (SLR). However, SLR+ SM gives better results and is broadly used to complement the SLR. The SLR is a form of a research study in which various research studies or papers are collected and critically analyzed systematically. An SLR intends to get an extensive and comprehensive summary of the relevant literature. Systematic mapping is a technique for gathering, organizing, and categorizing all available data on a particular topic and is beneficial at the start of the analysis as a brainstorming and scoping tool. Authors of [34] have looked at the difference between SLRs and SMs and concluded that in practice, only the quality assessment is performed differently in SLRs and SMs.

The systematic mapping process is iterative, with each phase building on the preceding one. Figure 1 depicts the procedure, which begins with a research question and concludes with a systematic map.

Figure 2 shows a flowchart demonstrating the procedure and decision-making process studied for inclusion and exclusion in this mapping review.

Manual search. To ensure that we find relevant articles, we conducted a manual search by reading the abstract of all papers published between 2019 and 2020 in the following journals and conferences: ACM Transactions on Computing Education[1], Computers & Education[2], IEEE Transactions on Education[11], Frontiers in Education (FIE) Conference[3], ICER[4], Computar science education[5], SIGCSE 2019[6], SIGCSE 2020[7], ISSEP 2019[8], ISSEP 2020[9] and Journal of Digital Learning in Teacher Education[10]. The purpose of the manual search was two-fold: (1) to fine-tune search terms and inclusion/exclusion criteria and (2) to create a reference set of papers to validate the results of the automatic search. To decide whether to exclude or include papers, we defined several criteria about report and study eligibility. More specifically, a paper should be included if published in English (RE1), if it has an abstract (RE2), if the publication is between 2010

[1]https://dl.acm.org/journal/toce
[2]https://www.journals.elsevier.com/computers-and-education
[3]https://ieeexplore.ieee.org/xpl/conhome/1000297/all-proceedings
[4]https://dl.acm.org/conference/icer
[5]https://www.tandfonline.com/toc/ncse20/current
[6]https://sigcse2019.sigcse.org/
[7]https://sigcse2020.sigcse.org/
[8]http://cyprusconferences.org/issep2019/
[9]https://issep2020.tlu.ee/
[10]https://www.tandfonline.com/toc/ujdl20/current

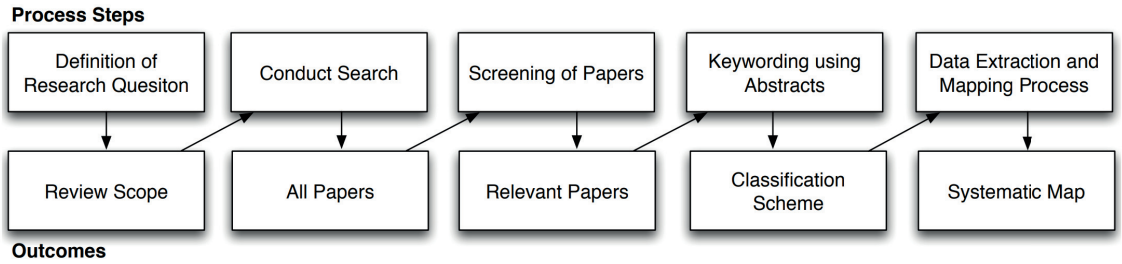

Fig. 1. The stages of the systematic mapping process [16]

and 2020 (RE3), and if it is a scientific paper and published in international conferences, peer-reviewed journals, or as book chapters (RE4). Further, a report should be included if it presents a study of K-12 teacher training focusing on the teaching of CS (SE1) and if it relates to the training of in-service teachers (SE2).

The purpose of the manual search was to confirm that the initial search keywords would find all the relevant papers. The results were then used to verify that the automated database search would find articles we were interested in. The research team used the identified papers (n=50) to refine the search criteria.

Data sources and literature search. Based on the results from the manual search and research questions, we defined the final search criteria to be used in the automatic database search as *(professional development OR teacher training OR teacher education) AND (computer science OR computational thinking OR programming OR informatics)*. This search query was divided into sub-queries Q1-Q12 as shown in Figure 2. The results from each sub-query were aggregated to a total number of 7251 papers. Further, we decided to perform the automatic search using the following databases: IEEE Xplore[11], ERIC[12], Science Direct[13], ACM Digital Library[14], Taylor & Francis Online[15], Wiley Online Library[16], and Springer[17].

The research team did the automatic search in several steps: 1) We had some initial training sessions to make sure that searches were run correctly and checked the results against the results of the manual search to assure that they were captured correctly. 2) A full automatic search was done in November 2020. 3) To include all papers in 2020, a new full search was done in July 2021 to capture papers that were published in the last part of 2020 and not accessible when the first searches were conducted. The screening process for the identified papers was unchanged. The numbers that are reported in Figure 2 are the total ones.

Screening of papers. The database search resulted in a total of n=7251 studies initially identified and reviewed based on their abstract to determine the appropriateness of inclusion. We imported the results into MS Excel and removed duplicates based on the titles. To simplify the screening process, we wrote an excel macro to traverse abstracts and highlight the keywords defined in the search criteria (professional development, teacher training, teacher education, CS, computational thinking, programming, informatics) with a different colour. This made the first step in the screening process more manageable as we could quickly eliminate the articles with no keywords. During the second step in the screening process, we checked for report eligibility criteria RE1-RE4 and study eligibility criteria SE1-SE2. To ensure all the group members have clear ideas about the selection criteria, two small sets of papers, including 50 and 100 articles, were selected and screened by each group member. The result showed that the paper selection match rate was 74%. The final results (n=206) were identified as relevant papers and transferred to NVivo for further analysis.

Method for Classification and Coding. After identifying relevant papers, we moved them to NVivo[18] and started the synthesis by analyzing the abstracts. We have used thematic analysis and the process described by [35] in this study. Based on the research questions, we defined an initial codebook with some high-level categories and codes. First, the three authors in the group familiarised with the data and research questions. We then coded the first twenty abstracts and used Fleiss' kappa [36] to measure inter-rater reliability for three coders[19]. Codes that emerged during the analysis were discussed and the codebook was updated. We repeated this process two times and reached inter-rater reliability of 78.3%, which we considered satisfactory. The remaining articles were distributed among the researchers to complete the coding. In addition to the coding of the abstracts, for each paper we recorded the number of authors as well as the type and geographical location of their affiliations.

The unit of coding varied from single words to short sentences when it was necessary to capture the context for further analysis. The resulting categories (based on paragraphs entities) were grouped/classified under each research question and presented in Section IV.

[11] https://ieeexplore.ieee.org/
[12]https://eric.ed.gov/
[13]https://www.sciencedirect.com/
[14]https://dl.acm.org/
[15]https://www.tandfonline.com/
[16]https://onlinelibrary.wiley.com/
[17]https://www.springer.com/

[18]https://www.alfasoft.com/en/products/statistics-and-analysis/nvivo.html
[19]https://www.statology.org/fleiss-kappa-excel/

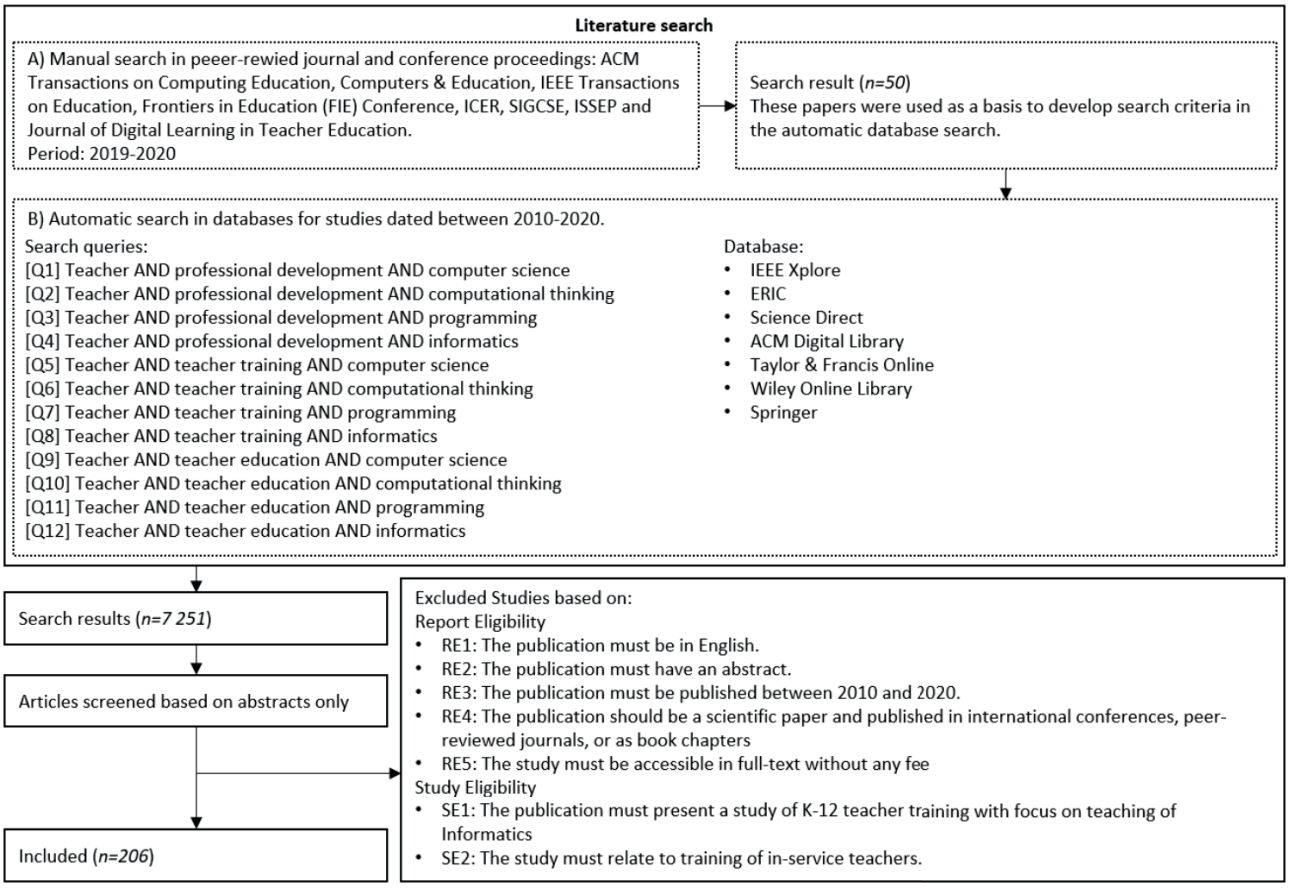

Fig. 2. Flowchart of search and screening process.

IV. Data Extraction and Mapping of Studies

The searching and screening resulted in 206 papers. All papers have an abstract between 58 and 387 words, with a median of 164 words. If we compare 2010 and 2020, the number of paper goes from 6 papers in 2010 to 42 in 2020. The number of authors also goes from 12 in 2010 to 153 in 2020, with more than 12 times. As a note of warning, this number is calculated by adding the number of authors for each paper without removing authors who have authored more than one paper during a year or in multiple years. So, it should be interpreted as the upper bound of the number of authors. Even considering duplicates, a total upper bound of 695 authors shows that there is now a community of considerable size around PD for in-service teachers of CS.

Research is conducted mainly by authors affiliated to universities, with universities being involved, either in partnership with others or standalone, in 94% of the papers. More in details, of the 206 papers, 146 papers have only authors affiliated to universities, 48 universities plus other types of affiliation, and 12 papers only authors not affiliated to universities.

Type of intervention (RQ.1). The first research question deals with the type of interventions reported in the literature. Identifying the type of intervention is a rather complicated aspect because different authors use different terms (for example, a workshop might describe short term interventions [37] to week-long residential interventions [4]). During the thematic analysis, we were able to identify 132 papers mentioning different types of PDs (See Figure 3). Forty-one talk about PD but do not specify the type of intervention in the abstract. Thirty-five studies mention workshops which seem to be the most common form of intervention. Thirty-one papers mention CS courses. Fourteen papers mention projects to support teachers' PD. Eleven papers refer to different forms of PD through cooperation among peers and Communities of Practice (CoP), for example, when the PD program includes establishing a CoP [38] or when there is some form of peer mentoring or group discussion [39].

A few papers explicitly refer to when/where the intervention takes place. Twelve papers refer to summer programs, including workshops, shorter courses, camps, or other activities. The abstracts do not specify under which conditions teachers participate and whether this is part of their working hours or done during their free time. Few interventions (5) are done in-school to help teachers develop their coding skills in the classroom, for example, [8].

When discussing the type of intervention, growing attention to sustainability is worth mentioning. 13 abstracts explicitly discuss sustainability, mostly from 2018 on. Sustainability is

discussed from two different perspectives. In some papers, sustainability refers to the need to develop PD programs that can be implemented with the available resources and local constraints, especially considering the need to train a high number of in-service teachers. In other papers, the focus is more on sustainability in terms of long-term learning processes and the changes brought by the PD.

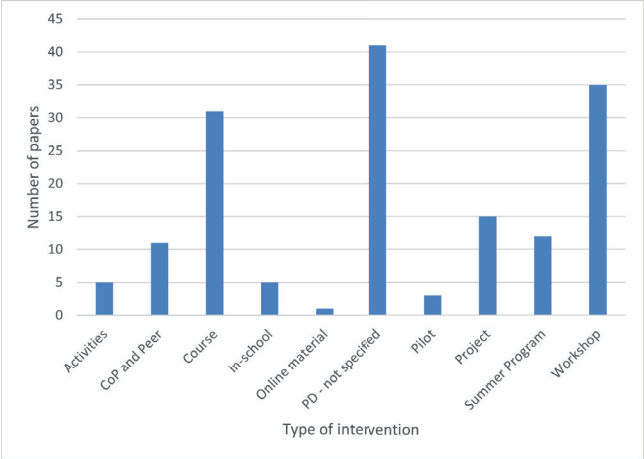

Fig. 3. Different types of interventions/studies identified during the coding process.

School level (RQ.2). Our second research question identifies which education level the intervention/research is intended to support. Since different terms are used in various countries [40], it was not easy to define clear boundaries for the intervention levels. Therefore, we report the results as reported in the literature. A total of 151 out of 206 have specified the school level for which the intervention is intended for. Some articles have specified several levels, for example primary and secondary. We have coded this as both primary and secondary levels. We identified the following school levels in the literature (The number in parenthesis shows the occurrence): K-12 (40), High-school 39), Primary (39), Secondary (28), Middle-school (25), K-9 (2), After-School Setting (1), Pre-collegiate (1), and Preschool (1). A minimum of 60% of the papers is about lower and upper secondary school (part of K-12 will also be in addition to this number). We can also see a noticeable number of studies in connection with primary schools. Suppose we assume a roughly equal distribution of secondary studies between middle- and high schools. In that case, we see that the focus has been approximately the same at primary (41) and middle school (39), while the number is higher for high school (53). We find few other types of school levels (preschools, after school settings, pre-collegiate), but the number of studies seems very low.

Classification of studies based on geographical area (RQ.3). For answering RQ.3, we identified in which country(ies) the research has been conducted and then categorized these within the meta-areas[20]. We have done this by looking at

[20]https://www.nationsonline.org/oneworld/continents.htm

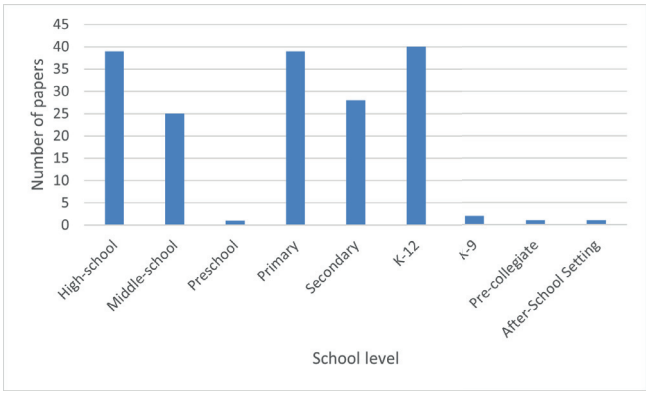

Fig. 4. Different school-levels for the interventions/studies identified during the coding process.

the locations of the authors' affiliations and explicit references to geographical areas in the abstracts. In total, 117 out of 206 papers have specified the place of intervention. Figure 5 shows this overview. We found no studies in African countries that met our eligibility criteria while analyzing the abstracts.

When we look at authors' affiliations, our data show that 92% of articles each have authors from the same area (North America 49%, Europe 22%, Asia 5%, the Middle East 5%, and Australia/New Zealand 5%).

Only 8% of papers have authors affiliated to institutions from more than one country, indicating international collaboration. The average number of authors for these papers is 4.6, higher than the overall average of authors per paper (3.4). The number of involved countries per paper varies between 2 and 5. All international collaborations are across meta-areas, except for four collaborations across European countries. Approximately half of the papers are before 2015 and a half after 2015. Therefore, international cooperation has not increased with the increase of papers in the area. This is a worrying element because one could expect an increased level of international cooperation as research grows and knowledge exchange becomes more beneficial. It should also be noted that authors from multiple countries do not necessarily imply an international study, i.e., studies of interventions across multiple countries. Of the 15 papers with international authors, only two are reporting international studies. One report on a national study but focuses on how this knowledge could be transferred to other countries.

Connection of intervention to curricula (RQ.4). A total of 104 out of 206 papers connects the type of intervention to different subjects in the curricula (See Figure 6). Two areas, in particular, stand out: CS (58) and STEM education (22). Other types of curricula identified are Algorithm (2), AI (1), Business (1), and English (1). Additionally, six papers referred to interventions connected to a cross-discipline curriculum. We could not see that the intervention is directly linked to a specific subject in about half of the studies. The trend shows an apparent increase from 2017 onwards in the type of intervention linked to a specific subject—about 80% of the

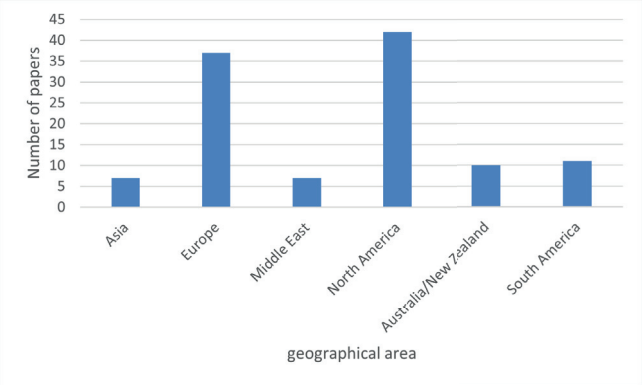

Fig. 5. Different geographical areas of the interventions/studies identified during the coding process.

cases we find in this period.

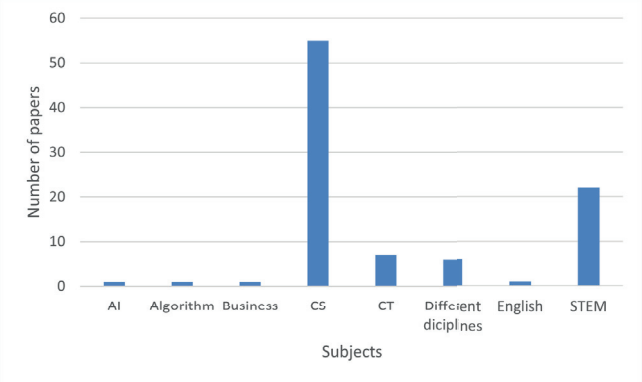

Fig. 6. Connection of intervention/study to curricula.

Learning objectives of the research (RQ.5). The research team was able to identify at least one learning objective of the PD intervention or program for 111 out of 206 papers (See Figure 7). Some of the identified objectives are rather broad, for example, improving knowledge of programming and programming didactics as in [41]. Other studies have a much narrower focus, such as integrating CS in specific subjects [42] or specific CS areas, like robotics or AI. One of the topics receiving great attention is Computations Thinking (CT), with 32 papers with CT as a learning objective. The focus on CT has gained momentum after 2017, with 90% of papers with CT as a goal published in 2017-2020. Another emerging theme that is worth mentioning is self-efficacy. Seventeen papers mention issues connected to understanding, measuring, or developing self-efficacy in teachers. All these papers are published after 2016. So, it seems there is an increased awareness in the community about the need to develop content and pedagogical knowledge and look at PD in the broader social context. This is also reflected in increased attention on equity and to need for teachers to learn to promote equitable access to CS education in schools. The terms equity/equitable are used explicitly in 13 papers (10 from 2018 on). However, the attention to access to CS education is more significant, with papers explicitly addressing issues connected to gender, special needs, race, and rural communities.

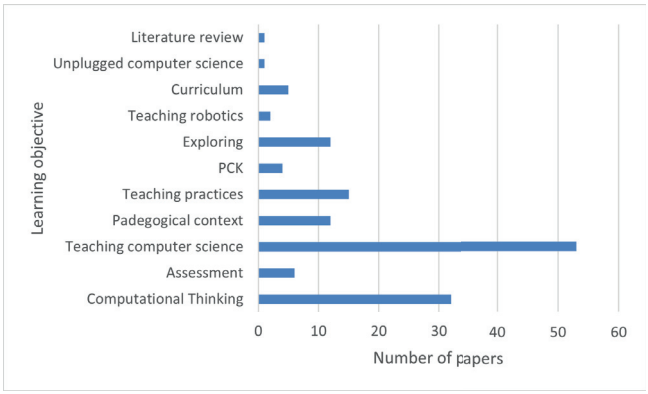

Fig. 7. Learning objectives of the research.

Phase of the research (RQ.6). While analyzing the abstracts, we identified four research phases for the presented studies: Inform Design, Design, Design & Evaluation, and Evaluation. Inform Design indicates studies that collect information with the aim of informing the design of PD programs, for example, the attitude of teachers and the challenges that they perceive to the introduction of CS in their teaching [43]. There is no description of a PD program for papers classified in this category, at least in the abstract. Studies under the category Design focus on the design of a PD intervention or program, while under Design & Evaluation, we classified papers that present the design together with its empirical evaluation, for example, "...providing empirical evidence of the design and evaluation of such a teacher development program..." [6]. Under Evaluation are the papers that present the empirical evaluation of an existing PD intervention or program. It should be noted that most of the studies categorized in Design & Evaluation or Evaluation are often part of an iterative process where the results of the evaluation are used to improve future phases of the PD implementation. For example, "...These initial findings suggest a pressing need for future research to look further into teacher CS self-esteem to inform..." [29]. No replication studies were identified. We were able to identify the research phase in 113 of 206 articles (See Figure 8). Most papers (42) fall under the category Design & Evaluation, 37 are under Inform design, 20 under Evaluation, and 16 under Design. In other words, about 69% of studies seem to be in Inform Design and Design and Evaluation. In general, this analysis presents a rich landscape of studies at different levels of maturity and promoting progress of the state of knowledge in different forms.

Size of the study (RQ.7). During the analysis, we classified the studies based on the number of involved teachers: small (< 25 teachers), medium ($25 \leq and \leq 100$) and large (> 100). We were able to identify the Size of the study in 59 of 206 articles. Most of these studies (29) are classified as large, 17 studies as small, and 12 as medium scale. It is important to underline that this should not be generalized as most of

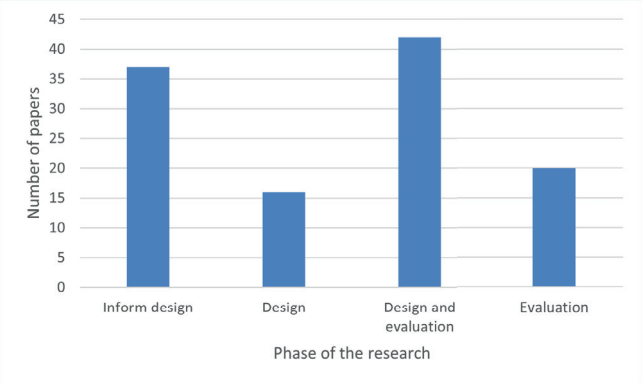

Fig. 8. Research phases for the presented studies.

the studies being large, since only around 25% of the papers specify the Size in the abstract. A closer look at the Size of interventions over time shows that approximately 69% of large studies are done after 2017. Due to the type of information extracted by the abstracts that mentioned the Size of the study, it is not possible to relate Size to other characteristics, for example, phase of the research or type of intervention.

Size should not be connected to the value of an intervention. Some interventions are small or medium scale, but they involve long-term cooperation with teachers, for example [8]. If we had classified the Size based on the number of hours of training, we might have got a different classification of the studies. The duration of the intervention, however, is seldom mentioned in the abstracts. The main result is that there are studies involving different numbers of teachers performed and accepted by the community.

Type of collaboration (RQ.8). Figure 9 shows an overview of type of collaboration over time. Only a few of the papers are single authored. The average number of authors for paper is 3.3. This is slightly higher than the number reported for ITiCSE papers (3.2, 2.7 if excluding working groups) [44]. It is also worth noticing that the average has been increasing. The average was 2.0 in 2010 and it has been over 3.0 since 2015 (3.6 in 2020). When we look at the affiliation of the authors, approximately 3/4 of the papers involve authors from only one organization (or departments from the same organization). Cross-sector cooperation, when looking at authors affiliation, is somehow limited, identifiable approximately in 25% of the papers. Most of the research is done by universities, but there are several other actors active in the area, including educational agencies, outreach programs, and IT companies. Schools are involved in teachers' Professional Development research only in six cases (3%), at least in terms of co-authorship. An in-depth analysis of the papers might lead to different results since collaboration is not always reflected in co-authorship. For example, in [45, p.992] the abstract states "The Louisiana Department of Education partnered with the …", but this is not evident in the authors' affiliation. The analysis of the full paper might help to understand the nature of the partnership. As a note, 8 papers explicitly mention the funding organizations: EU Erasmus+ (3), Google (2), Code.org (1), National Science Foundation (1), Naval Academy Foundation (1), NordForsk (1), TUBITAK (1). It is interesting to note that this includes both public and private funding. Funding of research is an important aspect that was addressed in [13]. However, the number of papers reporting it in the abstract is too limited to draw any conclusion. An analysis of the acknowledgement section of the papers might help to investigate this aspect further.

A. Emerging themes

Figure 10 shown the development of some patterns during the period 2010-2020 (RQ.9). The first trend to notice is that numerous studies have dramatically expanded over time. While 41 eligible research published between 2010 and 2014, there were 165 between 2015 and 2020. We find one paper in 2014 and one in 2015 that points to equity challenges and discusses inquiry- and equity-oriented approaches for broadening computing participation. This number was increased to 4 in 2018 and 8 in 2020, indicating more focus on the subject. Another area that seems to have changed significantly in recent years is CT. Between 2010 and 2016, we find ten papers discussing the inclusion of computational thinking into public schools, while we find 54 studies between 2017 and 2020. We also find that there is a greater focus on the development of self-efficacy in teachers. From 2018 onward, we have noticed a definite trend toward delivering teacher professional development to build engaging and long-term professional development models.

During the analysis, we noticed a drop in the number of eligible papers in 2019. Most themes seem to follow this drop except articles explicitly mentioning self-efficacy/attitude.

V. Discussion

In this section, based on the results presented in Section IV, we reflect on challenges and gaps identified in the literature.

Abstracts. The field misses a way to structure abstracts. Even for papers with long abstracts, there is considerable diversity in the type of information presented. This is a challenge shared with other areas of the computing education research community, as pointed out in [46, p.572], "…A paper's abstract has a valid purpose, … That purpose is to give readers a good idea of the content of the paper, to help them decide whether to take the trouble to read it.…". Good abstracts also help with developing high quality literature mappings to build overviews of the field. This is critical especially considering the relatively new and rapidly changing nature of research about PD of in-service teachers of CS.

Richness v. fragmentation. The papers that we analyzed present interventions of different types and sizes. This reflects a complex and rich landscape where PD must address very different needs, from providing basic competencies to integrate into various subjects to in-depth CS content and pedagogical knowledge training. Considering that CS has only been recently added in many countries or is in the planning, we expect that it will take some time before the

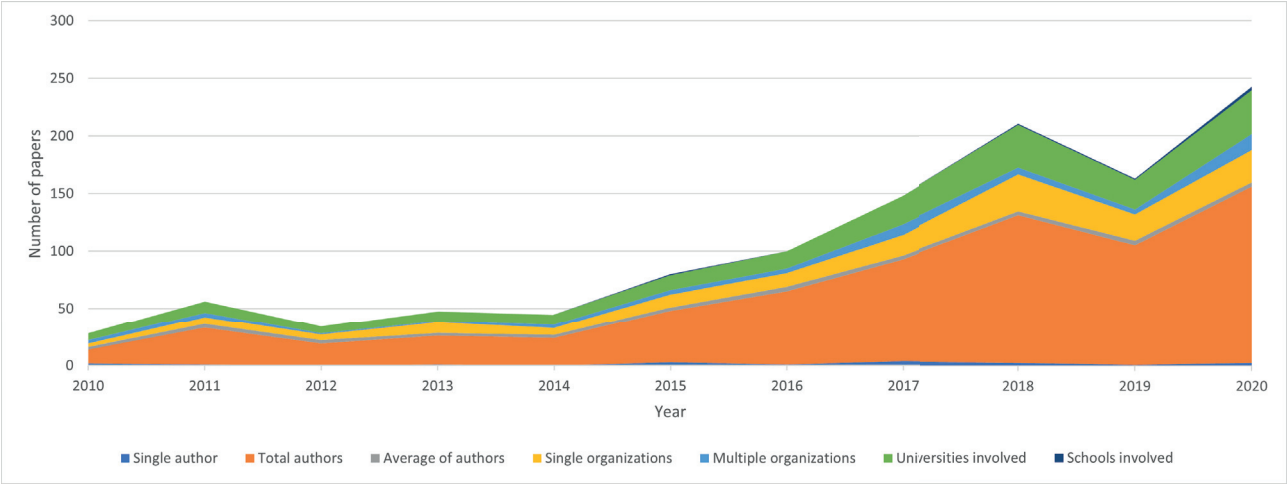

Fig. 9. Type of collaboration.

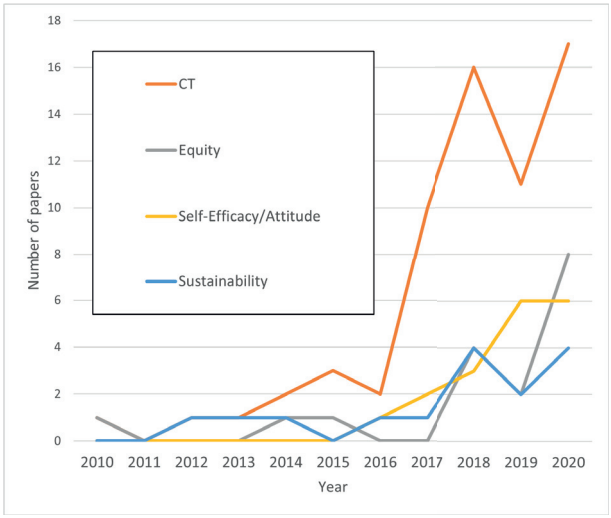

Fig. 10. Development of some patterns during the period 2010-2020

field can stabilize. Though this richness is positive, there is a challenge connected to the lack of shared terminology that might bring to a negative fragmentation of the field. This is partly common to all research in education. It is enough here to think about different ways to refer to educational levels in different countries [40]. Other challenges are more specific. For example, there is no ”standard” or structured way to describe the type of intervention. For instance, as discussed in Section IV,’ workshop’ is used for very different types of interventions in terms of form and duration. It is not clear what makes authors classify an intervention as a workshop rather than a course. Although there is no easy solution, it is essential to increase the awareness of the community on this challenge. Authors in writing their papers should be careful to make explicit the core characteristics of their work so that it is easier to understand. An effort of the community to define a shared terminology would also be helpful.

More inclusive research. As described in Section IV, there are only a few papers that involve authors affiliated to schools or the educational sector outside universities. We are aware that co-authorship is only one of the ways to show participation in research; still, this is a warning that might indicate that research is done on schools rather than with schools. This lack of cooperation was also reported in an earlier literature study, focusing on the United States [13]. So, this challenge seems to be more global and persisting in time. There are also geographical areas that are entirely missing, at least in the literature in English. A more global research community needs to be established to understand better challenges in regions with different cultural, economic, and educational backgrounds.

More international cooperation. Despite the field’s impressive growth, international cooperation, as described in Section IV, is still minimal. Only a few studies try to understand challenges across educational systems, either setting up joined studies or replicating interventions in different national settings. Multi-national studies in the educational sector are challenging but are critical to consolidating the knowledge in the area.

Limitation and threats. Literature mappings are, by nature, aiming at building a coarse mapping of a research field, focusing on breadth rather than depth. The strength of mapping is in capturing a broad overview, with overall trends and gaps, but it is impossible to investigate the different issues in depth. From this perspective, mapping is mainly a tool to help the research community to identify areas that might require more study. One of the threats of this work comes from the quality of the abstracts. As mentioned in the previous section, there is no established structure for papers in this research area, and there is a lack of shared terminology. In addition, 41% of the papers have less than 150 words, against a recommended length of 150–250 [47]. As a result, it was possible to extract only a limited amount of information for many abstracts, and they contributed marginally to answering the research questions.

In this perspective, the numbers that we provide should be interpreted cautiously. They are mainly intended to indicate trends rather than absolute numbers. Also, the mapping across studies has not been easy due to diverse educational systems and different terms. For example, in North America, the term K-12 is frequently used, while in Europe, pre-school, primary- and secondary school is used. A second threat that we want to mention is connected to selection biases. Two issues are worth mentioning. The first one is connected to the search terms that we have chosen. Though the initial manual search assured to capture a considerable number of relevant papers, some might have been missed. For example, [44] is a paper that could be included in terms of contents but was not captured with the searches. Though adding a snowballing step could have helped capture additional papers like this one, we decided to focus on the automatic searches for higher replicability of the process. The second, and most important, selection bias is that we considered only papers available in English. There might be a lot of local research that is not captured because published in national languages different than English. To address this limitation, we strongly encourage replicating similar mapping studies in other languages. In the analysis of the abstracts, we build on coding based on the research questions rather than a well-established categorization, like, for example, Simon´s classification system for computing education papers, used in [44]. However, to our knowledge, such classification does not exist for PD of in-service teachers of CS. So, we adopted a more open approach that helps to build the first overview of this rapidly changing research field.

VI. Conclusions

This paper presents a mapping of the literature published between 2010 and 2020 in English on Professional Development (PD) of in-service teachers of Computer Science (CS). Out of 7251 papers returned by automatic searches, 206 were selected as fulfilling the eligibility criteria. The abstracts of these papers have been analysed to build an overview of the research area. The paper presents the results of the analysis and identifies some gaps and areas requiring attention. Literature mappings have as a goal to build a general understanding of existing work in a domain area. In this perspective, our work can be seen as a starting point towards more in-depth studies in the area. As part of future work, we encourage replicating similar mapping studies in languages different from English. In addition, the community would benefit from a joint effort to define shared terminology and classification schemes that could be re-used across literature studies. Researchers could investigate several other issues by analysing different parts of papers, for example, research methods adopted in the field, funding models, schemes to support in-service teachers to participate in PD programs.

Acknowledgment

References

[1] P. Hubwieser, "The darmstadt model: a first step towards a research framework for computer science education in schools," in *International Conference on Informatics in Schools: Situation, Evolution, and Perspectives*. Springer, 2013, pp. 1–14.

[2] J. Vahrenhold, M. Caspersen, G. Berry, J. Gal-Ezer, M. Kölling, A. McGettrick, E. Nardelli, C. Pereira, and M. Westermeier, "Informatics education in europe: Are we all in the same boat?" 2017.

[3] C. Bellettini, V. Lonati, D. Malchiodi, M. Monga, and A. Morpurgo, "Informatics and computational thinking: A teacher professional development proposal based on social-constructivism," in *International Conference on Informatics in Schools: Situation, Evolution, and Perspectives*. Springer, 2018, pp. 194–205.

[4] J. Goode, S. R. Johnson, and K. Sundstrom, "Disrupting colorblind teacher education in computer science," *Professional Development in Education*, vol. 46, no. 2, pp. 354–367, 2020.

[5] Y. Dong, V. Cateté, N. Lytle, A. Isvik, T. Barnes, R. Jocius, J. Albert, D. Joshi, R. Robinson, and A. Andrews, "Infusing computing: Analyzing teacher programming products in k-12 computational thinking professional development," in *Proceedings of the 2019 ACM Conference on Innovation and Technology in Computer Science Education*, 2019, pp. 278–284.

[6] S.-C. Kong, M. Lai, and D. Sun, "Teacher development in computational thinking: Design and learning outcomes of programming concepts, practices and pedagogy," *Computers & Education*, vol. 151, p. 103872, 2020.

[7] J. Goode, K. Peterson, J. Malyn-Smith, and G. Chapman, "Online professional development for high school teachers: Building the capacity of a national cadre of teachers to broaden participation in computing," in *2019 Research on Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT)*. IEEE, 2019, pp. 1–5.

[8] H. E. Williams, S. Williams, and K. Kendall, "Cs in schools: Developing a sustainable coding programme in australian schools," in *Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education*, 2020, pp. 321–327.

[9] M. C. Stange, "First bytes: Preparing k-12 educators for cs inclusion," in *Proceedings of the 2020 ACM Conference on International Computing Education Research*, 2020, pp. 304–304.

[10] C. Fancsali, J. Mark, and W. Castillo, "Cs4all: An early look at teacher implementation in nyc," in *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*, 2020, pp. 1333–1333.

[11] A. Bollin and P. Micheuz, "Computational thinking on the way to a cultural technique," in *Open Conference on Computers in Education*. Springer, 2018, pp. 3–13.

[12] K. Falkner and R. Vivian, "A review of computer science resources for learning and teaching with k-12 computing curricula: An australian case study," *Computer Science Education*, vol. 25, no. 4, pp. 390–429, 2015.

[13] M. Menekse, "Computer science teacher professional development in the united states: a review of studies published between 2004 and 2014," *Computer Science Education*, vol. 25, no. 4, pp. 325–350, 2015.

[14] L. W. Wienhofen, B. M. Mathisen, and D. Roman, "Empirical big data research: a systematic literature mapping," *arXiv preprint arXiv:1509.03045*, 2015.

[15] K. Petersen, S. Vakkalanka, and L. Kuzniarz, "Guidelines for conducting systematic mapping studies in software engineering: An update," *Information and Software Technology*, vol. 64, pp. 1–18, 2015.

[16] K. Petersen, R. Feldt, S. Mujtaba, and M. Mattsson, "Systematic mapping studies in software engineering," in *12th International Conference on Evaluation and Assessment in Software Engineering (EASE) 12*, 2008, pp. 1–10.

[17] C. Mouza, A. Yadav, and A. Ottenbreit-Leftwich, "Developing computationally literate teachers: Current perspectives and future directions for teacher preparation in computing education," *Journal of Technology and Teacher Education*, vol. 26, no. 3, pp. 333–352, 2018.

[18] K.-. standards, "Standards — computer science teachers association," https://www.csteachers.org/page/standards, 2017, (Accessed on 07/17/2021).

[19] GOV.UK, "National curriculum - gov.uk," https://www.gov.uk/government/collections/national-curriculum, 2014, (Accessed on 07/17/2021).

[20] K. M. Rich, A. Yadav, and R. A. Larimore, "Teacher implementation profiles for integrating computational thinking into elementary mathematics and science instruction," *Education and Information Technologies*, vol. 25, no. 4, pp. 3161–3188, 2020.

[21] H. Bučková and J. Dostál, "Teachers' attitudes towards implementing coding at schools," in *2019 IEEE 15th International Scientific Conference on Informatics*. IEEE, 2019, pp. 000 335–000 340.

[22] A. Buss and R. Gamboa, "Teacher transformations in developing computational thinking: Gaming and robotics use in after-school settings," in *Emerging research, practice, and policy on computational thinking*. Springer, 2017, pp. 189–203.
[23] M. Borowczak and A. C. Burrows, "Ants go marching—integrating computer science into teacher professional development with netlogo," *Education Sciences*, vol. 9, no. 1, p. 66, 2019.
[24] A. Crowder, R. Dovi, and D. Naff, "Integrated stem education in virginia: Codeva elementary coaches academy," in *Integrated Approaches to STEM Education*. Springer, 2020, pp. 447–467.
[25] R. E. Ladner, A. Stefik, J. Naumann, and E. Peach, "Computer science principles for teachers of deaf students," in *2020 Research on Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT)*, vol. 1. IEEE, 2020, pp. 1–4.
[26] M. Coenraad, K. Mills, V. L. Byrne, and D. J. Ketelhut, "Supporting teachers to integrate computational thinking equitably," in *2020 Research on Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT)*, vol. 1. IEEE, 2020, pp. 1–2.
[27] K. Cabrera, P. Morreale, and J. J. Li, "Computer science+ education: an assessment of cs professional development," *Journal of Computing Sciences in Colleges*, vol. 33, no. 3, pp. 141–147, 2018.
[28] N. Monjelat, "Programming technologies for social inclusion: An experience in professional development with elementary teachers," in *2017 Twelfth Latin American Conference on Learning Technologies (LACLO)*. IEEE, 2017, pp. 1–8.
[29] R. Vivian, K. Quille, M. M. McGill, K. Falkner, S. Sentance, S. Barksdale, L. Busuttil, E. Cole, C. Liebe, and F. Maiorana, "An international pilot study of k-12 teachers' computer science self-esteem," in *Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education*, 2020, pp. 117–123.
[30] C. Fancsali, J. Mark, and L. A. DeLyser, "Nyc cs4all: An early look at teacher implementation in one districtwide initiative," in *2020 Research on Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT)*, vol. 1. IEEE, 2020, pp. 1–8.
[31] T. Toikkanen and T. Leinonen, "The code abc mooc: Experiences from a coding and computational thinking mooc for finnish primary school teachers," in *Emerging research, practice, and policy on computational thinking*. Springer, 2017, pp. 239–248.
[32] T. E. Reding, B. Dorn, N. Grandgenett, H. Siy, J. Youn, Q. Zhu, and C. Engelmann, "Identification of the emergent leaders within a cse professional development program," in *Proceedings of the 11th Workshop in Primary and Secondary Computing Education*, 2016, pp. 37–44.
[33] J. Flapan, J. J. Ryoo, and R. Hadad, "Building systemic capacity to scale and sustain equity in computer science through multi-stakeholder professional development," in *2020 Research on Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT)*, vol. 1. IEEE, 2020, pp. 1–8.
[34] B. Napoleão, K. R. Felizardo, É. F. de Souza, and N. L. Vijaykumar, "Practical similarities and differences between systematic literature reviews and systematic mappings: a tertiary study." in *SEKE*, 2017, pp. 85–90.
[35] V. Braun and V. Clarke, "Thematic analysis." 2012.
[36] R. Falotico and P. Quatto, "Fleiss' kappa statistic without paradoxes," *Quality & Quantity*, vol. 49, no. 2, pp. 463–470, 2015.
[37] J. Simmonds, F. J. Gutierrez, C. Casanova, C. Sotomayor, and N. Hitschfeld, "A teacher workshop for introducing computational thinking in rural and vulnerable environments," in *Proceedings of the 50th ACM Technical Symposium on Computer Science Education*, 2019, pp. 1143–1149.
[38] A. Joshi, A. Jain, E. Covelli, J.-h. Yeh, and T. Andersen, "A sustainable model for high-school teacher preparation in computer science," in *2019 IEEE Frontiers in Education Conference (FIE)*. IEEE, 2019, pp. 1–9.
[39] D. J. Ketelhut, K. Mills, E. Hestness, L. Cabrera, J. Plane, and J. R. McGinnis, "Teacher change following a professional development experience in integrating computational thinking into elementary science," *Journal of science education and technology*, vol. 29, no. 1, pp. 174–188, 2020.
[40] Wikipedia, "Educational stage," 2 2015, (Accessed on 07/15/2021). [Online]. Available: https://en.wikipedia.org/wiki/Educational_stage
[41] G. Ahmed, J. Nouri, L. Zhang, and E. Norén, "Didactic methods of integrating programming in mathematics in primary school: Findings from a swedish national project," in *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*, 2020, pp. 261–267.
[42] L. Ilona-Elefteryja, M. Meletiou-Mavrotheris, and K. Katzis, "A teacher professional development program on teaching stem-related topics using augmented reality in secondary education," in *Emerging Technologies and Pedagogies in the Curriculum*. Springer, 2020, pp. 113–126.
[43] P. Vinnervik, "Implementing programming in school mathematics and technology: teachers' intrinsic and extrinsic challenges," *International journal of technology and design education*, pp. 1–30, 2020.
[44] J. Sheard, "Twenty-four years of iticse papers," in *Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education*, 2020, pp. 5–11.
[45] F. Alegre, J. Underwoood, J. Moreno, and M. Alegre, "Introduction to computational thinking: a new high school curriculum using codeworld," in *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*, 2020, pp. 992–998.
[46] J. Sorva, "How concrete should an abstract be?" in *Proceedings of the 26th ACM Conference on Innovation and Technology in Computer Science Education V. 1*, 2021, pp. 568–574.
[47] S. F. Gambescia *et al.*, "A brief on writing a successful abstract," *Education for Health*, vol. 26, no. 2, p. 122, 2013.