Assessment of learning in child–computer interaction research: A semi-systematic literature review

Citation for the original published paper (version of record):
Elif Baykal, G., Eriksson, E., Torgersson, O. (2023). Assessment of learning in child–computer interaction research: A semi-systematic literature review. International Journal of Child-Computer Interaction, 36.
http://dx.doi.org/10.1016/j.ijcci.2023.100578

(article starts on next page)

International Journal of Child-Computer Interaction 36 (2023) 100578

Contents lists available at ScienceDirect

International Journal of Child-Computer Interaction

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

Review article

Assessment of learning in child–computer interaction research: A semi-systematic literature review

Gökçe Elif Baykal [a,*], Eva Eriksson [b], Olof Torgersson [c]

[a] Özyeğin University, Turkey
[b] Aarhus University, Denmark
[c] University of Gothenburg, Sweden

ARTICLE INFO

Article history:
Received 3 September 2022
Received in revised form 28 February 2023
Accepted 28 February 2023
Available online 16 March 2023

Keywords:
Child–computer interaction
HCI
Learning
CCI

ABSTRACT

In this paper, we investigate and map out how learning is assessed in Child–computer interaction (CCI) research. We have conducted a semi-systematic literature review in the CCI community's leading venues: the Interaction Design and Children (IDC) conference and the International Journal of Child–Computer Interaction (IJCCI). This eventually led to 30 publications that use the word stem 'learn*' in title, abstract and keywords being included in the corpus. Based on our analysis of these publications, the results demonstrate that there are three main strands of research approaches, namely quantitative, qualitative and mixed-methods, some of which are design-based. The case studies taking a qualitative approach dominate the field whereas the mixed-methods approach remains low in number. Furthermore, the findings showed that basic characteristics of research design and approaches to the assessment of learning are rarely defined, and that assessment of learning is scarcely operationalized. This affects the methodological rigor and possibility of understanding causality of technology interaction in children's learning. It was also found that only a limited number of works include assessment of learning regarding transfer of learning and controlled groups. The main findings from this review describe the current state-of-the art and address the gaps in CCI research in presenting evidence for learning in children as a desired impact. We conclude with suggestions for future avenues for the assessment of learning in CCI.

Contents

* Corresponding author.
E-mail address: elif.baykal@ozyegin.edu.tr (G.E. Baykal).

https://doi.org/10.1016/j.ijcci.2023.100578

1. Introduction

Child–computer interaction (CCI) is a multidisciplinary field that concerns the phenomena surrounding the interaction between children and computational and communication technologies (Read et al., 2013) and has always had a strong connection to learning and education (Giannakos, Horn, Read and Markopoulos, 2020; Giannakos, Papamitsiou, Markopoulos, Read and Hourcade, 2020). Starting in the 1960's (e.g. Ackermann, 1991; Papert, 1980), CCI research work stem largely from the field of Human–Computer Interaction (HCI). However, as a multidisciplinary research community, CCI is directly connected with several research areas such as psychology, learning sciences, interaction design, engineering, computer science, and media studies. This means that CCI draws on a wide range of research methods and practices from various fields, and within each discipline, there are different research traditions, which requires that we approach research based on different criteria (Eriksson, Baykal, & Torgersson, 2022; Hourcade, 2007). As a community, we need to appropriate and adapt those research methods, by contextualizing them to CCI and making them relevant to CCI challenges (Giannakos, Horn et al., 2020). Today, learning in education and beyond are still driving forces in CCI (Giannakos, Papamitsiou et al., 2020), and the community is still reinforcing natural ties with both high-quality HCI and learning conferences (Giannakos, Horn et al., 2020).

As CCI as a field concerns the phenomena surrounding the interaction between children and computational technologies (Read et al., 2013), we typically evaluate some technology, based on some criteria (e.g. usability, effectiveness, satisfaction, usefulness, efficiency, learnability, fun, accessibility, safety), and with a specified goal (Markopoulos, Read, MacFarlane, & Hoysniemi, 2008). Particularly tricky is the need to make a distinction between goals of product use and goals for evaluation. For example, even though the main purpose of a product X might be "learning new skill A" the evaluation of the product might focus on the usability of the user interface rather than on measuring the learning of skill A (Markopoulos et al., 2008). It is also possible to include aspects of learning as criteria for the evaluation of some technology, where the goal of the evaluation is the assessment of learning (Cukurova & Luckin, 2018). There is a distinction here, where the technology is being *evaluated*, while the *assessment* rather concerns determination of students' status in terms of learning (Popham, 2008). However, in a recent literature review on the role of learning theory in CCI, Eriksson et al. found that criteria for learning are rarely addressed in CCI research (Eriksson et al., 2022).

In this paper we will build on and extend the mentioned findings by mapping out and investigating *how* learning is assessed in CCI research. We apply the same method as in Eriksson et al., namely a semi-systematic literature review of research papers from the Interaction Design and Children Conference Proceedings (IDC) and International Journal of Child–Computer Interaction (IJCCI). Based on this corpus, we outline what aspects of learning are assessed in CCI, and what research approaches, designs, methods and variables are used for the assessment of learning in CCI. To our knowledge, such a review has not previously been conducted, but is timely as the increased innovation and use of emerging technologies in education are leading to the development of many new learning opportunities (Council, 2000), and thereby also for CCI as a field to develop further (Eriksson et al., 2022).

2. Background

Learning is a complex phenomenon, it ranges from learning motor skills to learning to become a professional, and there is no one definition of learning that is universally accepted by theorists, researchers, and practitioners (Schunk, 2012; Shuell, 1986). Although there is a disagreement about the precise nature of learning, and regardless of the theoretical stance one takes on learning, be it behavioral, cognitive, constructivist, situated, sociomaterial, or some other theoretical orientation, they all share a fundamental understanding that learning involves "a change" in behavior or cognitive process (Gajda, Karwowski, & Beghetto, 2017). According to Schunk, there are three criteria for learning: (i) learning involves change, (ii) learning endures over time, and (iii) learning occurs through experience (Schunk, 2012). This means that we may define learning as a change in probability of response, but we must also specify the conditions under which it comes about.

Learning is complex and may have an impact on our interventions and for how we plan, execute, and evaluate research in educational situations and contexts. Therefore, we need to have a basic understanding for the assessment of learning, which involves "a formal attempt to determine students' status with respect to educational variables of interest" (Popham, 2008, p. 6). Typically, learning is assessed based on what people say, write, or do, using different types of assessment methods. However, as learning involves "a change" in behavior, or a changed capacity to behave in a given fashion, we must also take into consideration that it is not uncommon for people to learn skills, knowledge, beliefs, or behaviors without demonstrating them at the time learning occurs (Schunk, 2012).

Researchers and practitioners who want to know whether learning has occurred, may use procedures other than testing that provide evidence of student learning, see Table 1 adapted

Table 1
Methods of assessing learning, from Schunk (2012) p.15.

Category	Definition
Direct observations	Instances of behavior that demonstrate learning
Written responses	Written performances on tests, quizzes, homework, papers, and projects
Oral responses	Verbalized questions, comments, and responses during learning
Ratings by others	Observers' judgments of learners on attributes indicative of learning
Self-reports	People's judgments of themselves through: Questionnaires, Interviews, Stimulated recalls, Think-alouds, Dialogues

from Schunk (2012). Second, students' skills in content areas often are the learning outcome assessed, but researchers and practitioners may also be interested in other forms of learning. For example, they may want to know whether students have learned new attitudes or self-regulation strategies or whether students' interests, values, self-efficacy, and motivation have changed as a result of content learning (Schunk, 2012).

It is common for educational innovations, such as new learning resources, to be evaluated by an experiment where learning gains or other desired outcomes are compared between an experimental condition involving the innovative experimental 'treatment' and some comparison condition where the treatment being evaluated is absent (Taber, 2019). However, a wide variety of legitimate scientific designs are available for educational research (National Academies Press, 2002), which means that experimental designs are very suitable for some educational studies, but are not indicated for others, and that experimental research can be productively complemented by other forms of enquiry (Taber, 2019). For instance, Collins, Joseph, and Bielaczyc compared how design studies differ from laboratory experiments in the field of education (Collins et al., 2004). Design-based research deals with real world situations that contain limitations, complexities, and dynamics, while laboratory experiments are conducted in the laboratory without significant interruption from other variables.

According to Schunk, there are three research paradigms in learning: *Correlational*, which examines relations between variables, *Experimental*, where one or more variables are altered and effects on other variables are assessed, and *Qualitative*, which is concerned with description of events and interpretation of meanings. However, correlational findings often suggest directions for further research, and are often not what is found in publications, which is why we have chosen not to code for that in this review. Instead, we have coded for qualitative, quantitative, mixed-methods, and also design-based research approaches in which learning has been assessed (American Psychological Association, 2022a; Collins et al., 2004; Creswell & Creswell, 2017; Schunk, 2012). The four research approaches should not be viewed as rigid, distinct categories, opposites, or dichotomies. Instead, they represent different ends on a continuum (Creswell & Creswell, 2017), and a study tends to be more qualitative than quantitative or vice versa, etc. We further code for what specific research design has been applied (American Psychological Association, 2022a; Creswell & Creswell, 2017).

There are many different discourses when it comes to evaluating technologies in educational contexts. For instance, learning technology research focuses on the design, development, and/or use of technologies that support learning, whereas CCI research focuses on the design, development, and/or use of technologies that support children's lives (with a heavy emphasis on learning) (Giannakos, 2022). The main difference is the focus, namely that CCI adopts a child-centered perspective, and learning technologies a learner-centered perspective. Although the two fields have many similarities in method, this difference in focus leads to methodological differences. In CCI, methodological approaches have been designed that embrace the fact that children are central participants and not just users of the technology (Frauenberger, Good, Fitzpatrick, & Iversen, 2015; Iversen, Smith, & Dindler, 2017). Also, due to the focus on the child in CCI research, in many cases it is not possible to collect data directly from the child, but one must rather rely on proxy data from the children's support sphere (e.g., surveys or interactions with teachers or parents) (Giannakos, 2022).

However, there are many related and interesting methodological discussions in fields such as learning technologies that CCI can be inspired by in terms of evaluating technology in educational contexts. For instance, in psychological research, based on a review on how to do research where the aim is to use some intervention to improve cognitive skills, Shawn Green et al. call for improving methodological standards in behavioral interventions for cognitive enhancement (Shawn Green et al., 2019). Also, in educational research there is the Clark/Kozma debate concerning that the medium used for teaching does not matter, that it is all about the instructional design, and that using some technology does not, in principle, add anything that cannot be achieved without the technology. Becker argues against this in the area of educational games and claims that the game design and the instructional design taken together matter (Becker, 2010). In the EdTech field, there is the tension between experimental research and iterative design-based research, where authors advocate for a broader view on assessment as technology evolves so fast (Cukurova & Luckin, 2018; Cukurova, Luckin, & Clark-Wilson, 2019). Although we acknowledge the potential value and inspiration for the CCI research community to find in these related fields, in this review we will keep the focus to CCI.

3. Method

Inspired by Eriksson et al. (2022) and Lyle, Korsgaard, and Bødker (2020), we are doing a semi-systematic literature review (Snyder, 2019; Wong, Greenhalgh, Westhorp, Buckingham, & Pawson, 2013). A semi-systematic review is suitable for topics that have been differently conceptualized and studied by different groups of researchers and where it is not possible to review the whole field, but rather make an overview of the topic (Snyder, 2019; Wong et al., 2013). The semi-systematic review was chosen for this paper as CCI is a multidisciplinary field (Giannakos, Horn et al., 2020), and because in order to find a representative sample we limited the search to two venues and to publications that make use of the search term learn* in title, abstract, and keywords. Accordingly, the review makes no claim to cover all existing research in CCI from every angle possible, but the reviewed papers have been systematically coded to identify and provide an overview of potentially relevant research approaches to assess learning in CCI.

For the IDC conference we used the following search query in the ACM digital library:

- Title:(learn*) AND Abstract:(learn*) AND Keyword:(learn*) "filter": Article Type: Research Article, Conference Collections: IDC: Interaction Design and Children, ACM Content: DL.

For IJCCI, we used the following search query in Scopus:

- (TITLE (learn*) AND ABS (learn*) AND KEY (learn*) AND ISSN (22128689)).

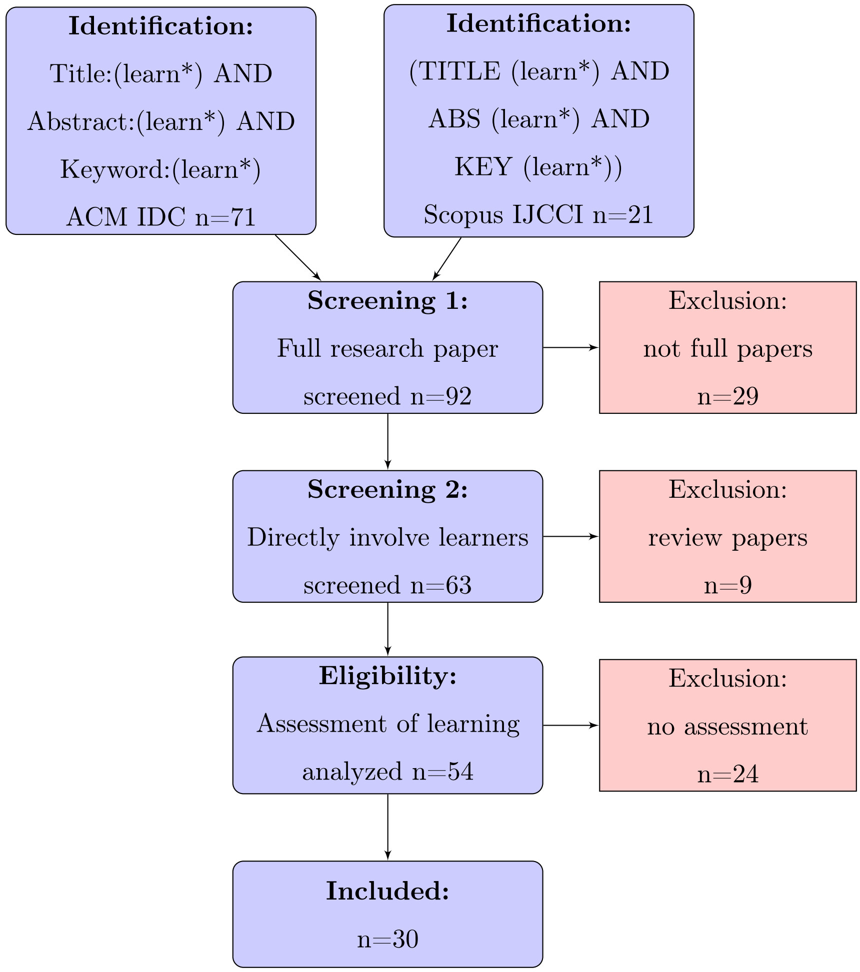

Fig. 1. Flow diagram of semi-structured literature review (based on Moher, 2009).

We had no further limitation in terms of years or age of participants.

The search, performed in October 2021, resulted in 92 publications, 71 from IDC, and 21 from IJCCI. The criteria for inclusion were that (1) the publication was a full research paper (which excluded 29 IDC publications such as late breaking work, doctorial consortium, demos, etc.), (2) learners are directly involved in the assessment (which excluded nine interview studies and review papers), and finally (3) the publication assesses learning (which resulted in 24 excluded papers). The decision for including full papers only is due to concerns that there would otherwise not be enough space to elaborate the methodological underpinning of the assessment. The reasoning for the second criteria was that when there is no direct involvement of learners, there cannot be any assessment of learning. This was the case in the excluded literature review papers (Baykal, Alaca, Yantaç, & Göksun, 2018; Flynn, Kleinknecht, Ricker, & Blumberg, 2021; Manches & Price, 2011; Oranç & Küntay, 2019; Roskos, Brueck, & Lenhart, 2017; Troseth & Strouse, 2017), and interview papers (Bunting, af Segerstad, & Barendregt, 2021; Einarsson & Hertzum, 2020; Tisza et al., 2020).

For the third inclusion criteria, we performed a round of more detailed coding regarding the assessment of learning. Two of the authors coded all 54 papers individually. In cases where there was no consensus, the third author took the decision. In this second more detailed round of coding, we found that many papers do not mention any type of assessment of learning and they were therefore excluded. Other criteria for exclusion were that the authors explicitly state that they have not studied learning (Fan & Antle, 2020; Li, van der Spek, Yu, Hu, & Feijs, 2020; Tisza & Markopoulos, 2021; Yip et al., 2014), and also when studying behavior in a learning situation using a technology, but without presenting the results e.g. Wyeth and MacColl (2010). This led to 24 excluded papers. The process is summarized in Fig. 1. In the end, a total of 30 papers were included in the final corpus, 24 from IDC and 6 from IJCCI, covering the years 2009–2021, see Table 2.

We recognize that our search strategy may have failed to generate a set of articles without a biased representation. The review is limited by the search query itself, as the term learn* is not covering all relevant CCI research. We further acknowledge that the review is limited by only including two publication venues, although CCI have many other publication venues. As such, we make no claims to cover the whole CCI field, but rather a sample. However, the included corpus has been systematically analyzed.

3.1. Analysis of data

All three authors manually coded all the publications separately and thoroughly. In case of discrepancy, the authors resolved it through discussion. This process of open and iterative coding was repeated by all authors more than twice for all 30 papers to increase consistency and to ensure that all categories were saturated, meaning that continuing analysis would have diminished returns in the new categories.

Table 2
Overview of the final corpus with 30 included publications — 24 from IDC and 6 from IJCCI.

Year	IDC	IJCCI
2009	Moher (2009)	
2010	Garzotto and Bordogna (2010)	
2011	Antle, Wise, and Nielsen (2011) Rick, Marshall, and Yuill (2011) Lamberty, Adams, Biatek, Froiland, and Lapham (2011)	
2012	Malinverni, Silva, and Parés (2012)	
2013	Shimoda, White, Borge, and Frederiksen (2013) Bartoli, Corradi, Garzotto, and Valoriani (2013)	
2015	Apostolellis and Bowman (2015) Chase and Abrahamson (2015)	
2016	Pantic, Fields, and Quirke (2016) Guo et al. (2016) Bell and Davis (2016) Johnson, Shum, Rogers, and Marquardt (2016a) Apostolellis and Bowman (2016a)	Tan, Goh, Ang, and Huan (2016)
2017	Keifert et al. (2017)	Westlund et al. (2017)
2019	Michaelis and Mutlu (2019) Beheshti, Borgos-Rodriguez, and Piper (2019)	
2020	Wan, Zhou, Ye, Mortensen, and Bai (2020) Zimmermann-Niefield, Polson, Moreno, and Shapiro (2020) Ruan et al. (2020)	Vartiainen, Tedre, and Valtonen (2020)
2021	Ho, Cagiltay, White, Hubbard, and Mutlu (2021) Tseng et al. (2021)	Tisza and Markopoulos (2021) Lee-Cultura, Sharma, and Giannakos (2021) Flores-Gallegos, Rodríguez-Leis, and Fernández (2021) Vartiainen et al. (2021)

3.1.1. Descriptive information

Regarding descriptive information, we have coded for the *country* and *context* where the research was conducted, *age* and *type* of learners, and *type of technology* involved in the reported research. For *type of learning*, we make a distinction between *Formal Learning*, *Non-formal Learning* and *Informal Learning* (COE, 0000; Eshach, 2007; OECD & Development, 0000; Tisza et al., 2019).

We coded for four learning domains as defined in the Developmentally Situated Design (DSD) cards (Bekker & Antle, 2011): *Cognitive* - reasoning, problem solving and logic skills (e.g., mathematics and language learning), *Emotional* - emotion expression, emotion regulation, emotion understanding, *Social* - self-esteem, perspective taking, friendships, and *Physical* - locomotion, manipulation, and stability.

3.1.2. Learning criteria

As presented previously, we have adopted the three criteria for learning as defined by Schunk (2012). To estimate the first criteria, (i) learning involves change, one should examine the difference between the current and the desired state of a learner. Administrating pre- and post-tests is the most common method to investigate any possible change occurring before and after an intervention or learning experience. The second criteria, (ii) endures over time, could be investigated as to whether learning endures over time by looking at a delayed post-test, or if the knowledge is carried to another context by looking at transfer of learning. Finally, (iii) occurs through experience, whether a particular experience has an impact on learning could be assessed through a controlled experiment by comparing a different but equivalent sample where the experience does not exist.

3.1.3. Learning outcome and dependent variable

To understand the assessment of learning, we also analyzed the *learning outcome*, and *dependent variable(s)* (American Psychological Association, 2022b) or the predicted outcome variable being investigated in each publication. The rationale behind this is to grasp the effects for which CCI researchers sought to find the cause of learning or expected to find a change in outcome of a learning experience. The *dependent variable* is a response variable or output that is measured in a study and is expected to change as a result of the researcher's manipulation of the independent variable in experimental research (American Psychological Association, 2022b; Boudon, 1969; Feldman & Garrison, 1993; Tuckman, 2012). The dependent variable is dependent on the actions of the participants or subjects — the people taking part (or recruited) in the experiment. However, studies in the CCI field are not necessarily conducted in an experimental research design fashion. Thus, we coded any predicted outcome variables mentioned in the studies as well as *dependent variables* being measured – if any – as the criteria or predictor of a learning gain in the reviewed studies. We also coded the methods, tools and techniques used for measuring learning along with whether pre- and post-tests were administered, and if a control group and/or transfer of learning (Barnett & Ceci, 2002; Kolodner, Gray, & Fasse, 2003; Schwartz, 2005) were examined or not.

3.1.4. Research approach and research design

We have coded for *qualitative research, quantitative research* including experimental research, *mixed-methods research*, and also *design-based research* as approaches in which learning has been assessed (American Psychological Association, 2022a; Brown, 1992b; Collins et al., 2004; Creswell & Creswell, 2017; Plomp, 2013; Schunk, 2012). For all papers, especially for quantitative studies, we have also coded for if a *hypothesis* was used and as to whether experiments include true experiments, with the random assignment of subjects to treatment conditions or quasi-experiments that use nonrandomized assignments (Creswell & Creswell, 2017; Keppel & Wickens, 2004). We also checked whether a statistical power analysis was conducted to estimate the target sample size which is important when a study aims to detect the relation between specific variables and measure the probability of correctly rejecting the null hypothesis (Cohen, 2013; Mayers, 2013). We further coded for what specific research design has been applied. To analyze the characteristics of research design, we made an a priori classification based on the 11 definitions, e.g. *One-Shot Case Study, One-Group Pretest-Posttest Design, Pretest–Posttest Control-Group Design, Solomon Four-Group Design,* etc. found in Creswell and Creswell pp. 231–41. These categories served us as a template for analyzing the methodological procedures carried out for assessment of learning. We have also coded the corpus according to if the research is conducted in a *laboratory* or *field* context (Schunk, 2012).

4. Results

4.1. Descriptive information

When coding the corpus, we see that 29 of the 30 papers are in the cognitive learning domain. One paper is in the social domain (Tan et al., 2016), one paper is a combination of the cognitive and physical learning domains (Lee-Cultura et al., 2021), two papers are a combination of cognitive and emotional (Bartoli et al., 2013; Vartiainen et al., 2020), and six papers are a combination of the cognitive and social domain (Apostolellis & Bowman, 2016a; Bell & Davis, 2016; Johnson et al., 2016a; Keifert et al., 2017; Lamberty et al., 2011; Westlund et al., 2017).

The learning domain is strongly connected to the subject of study, what the children are learning about. The largest category is science with eight papers (Beheshti et al., 2019; Bell & Davis, 2016; Guo et al., 2016; Keifert et al., 2017; Malinverni et al., 2012; Michaelis & Mutlu, 2019; Moher, 2009; Shimoda et al., 2013), math six papers (Chase & Abrahamson, 2015; Ho et al., 2021; Lamberty et al., 2011; Lee-Cultura et al., 2021; Rick et al., 2011; Ruan et al., 2020), technology eight papers, five of which focus on machine learning (ML) (Tseng et al., 2021; Vartiainen et al., 2020, 2021; Wan et al., 2020; Zimmermann-Niefield et al., 2020) and three papers on programming (Johnson et al., 2016a; Pantic et al., 2016; Tisza & Markopoulos, 2021), language three papers (Flores-Gallegos et al., 2021; Garzotto & Bordogna, 2010; Westlund et al., 2017), social skills (Tan et al., 2016), other more narrow subjects such as sustainability (Antle et al., 2011), olive oil production (Apostolellis & Bowman, 2015, 2016a), and attention skills (Bartoli et al., 2013).

The most common form of learning is *informal learning*, as found in 15 papers, followed by *formal learning* found in 14 papers, and *non-formal learning* in only one paper.

Most of the studies have been conducted in North America (USA 17, Canada 1) and Europe (10), which is not surprising as the IDC conference mainly takes place in USA or Europe. There is further one paper from Mexico, and one from Singapore.

The participants involved in the studies, the learners, are most often children age 10–15 years (17 papers), age 5–10 years (8 papers), age 0–5 years (2 papers), 15–18 (2 papers), or not specified (one paper). In terms of type of learner, studies most commonly identified learners in terms of age or school grade except for five papers; one paper included children with reading disabilities (Flores-Gallegos et al., 2021), one with autistic children (Bartoli et al., 2013), one partly with low socio-economic background (Apostolellis & Bowman, 2016b), one with disabled children (Garzotto & Bordogna, 2010), and one paper on gifted children (Lamberty et al., 2011).

The technologies reported on in the studies varies and are often combined, but we can see some clusters, such as games (Apostolellis & Bowman, 2015, 2016a; Bartoli et al., 2013; Ho et al., 2021; Lee-Cultura et al., 2021; Malinverni et al., 2012; Tan et al., 2016), block programming languages (Guo et al., 2016; Pantic et al., 2016; Tisza & Markopoulos, 2021; Zimmermann-Niefield et al., 2020), machine learning (Vartiainen et al., 2020, 2021; Wan et al., 2020; Zimmermann-Niefield et al., 2020), robots (Ho et al., 2021; Michaelis & Mutlu, 2019; Westlund et al., 2017), extended reality (VR or MR) (Flores-Gallegos et al., 2021; Keifert et al., 2017), and maker technologies including micro:bit (Johnson, Shum, Rogers, & Marquardt, 2016b; Tisza & Markopoulos, 2021; Tseng et al., 2021).

In the analysis, we see that in 7 out of 30 papers the research was conducted in a lab (Flores-Gallegos et al., 2021; Lee-Cultura et al., 2021; Malinverni et al., 2012; Michaelis & Mutlu, 2019; Ruan et al., 2020; Wan et al., 2020; Zimmermann-Niefield et al., 2020). This means that most of the studies take place in the field, such as museums, schools, therapy center, summer camp, etc.

4.2. Research approaches in assessing learning

4.2.1. Quantitative research approach

We found 10 experimental studies (Apostolellis & Bowman, 2015, 2016a; Chase & Abrahamson, 2015; Flores-Gallegos et al., 2021; Johnson et al., 2016a; Malinverni et al., 2012; Michaelis & Mutlu, 2019; Ruan et al., 2020; Tan et al., 2016; Westlund et al., 2017), in which some qualitative methods were also partly used, and one non-experimental study based solely on a quantitative approach (Tisza & Markopoulos, 2021) to assess the learning outcome in their study. All of these 11 studies presented a hypothesis related to learning in children and tested the hypothesized effect on learning in the study. In nine of these 11 papers, a hypothesis about learning was clearly introduced earlier than the Methods section in the paper and tested in the study, and in two papers (Flores-Gallegos et al., 2021; Westlund et al., 2017) the hypothesis was vaguely presented in the Discussion section.

In addition to the 11 papers, there were also four papers (one case study Guo et al., 2016, one field study Bartoli et al., 2013 and two design-oriented research Antle et al., 2011; Beheshti et al., 2019) that partly utilized quantitative research approaches in reporting the learning outcomes. However, none of these four studies presented a hypothesis on learning.

When it comes to estimating the sample size to assess the hypothesized effect on learning, only three papers (Apostolellis & Bowman, 2015, 2016a; Tan et al., 2016) reported calculating the statistical power analysis in their research planning. In fact, two of these papers belong to the same broader project and are thus based on the same sampling procedure. Moreover, Michaelis and Mutlu (2019) mentioned the lack of doing a power analysis as a limitation in the study and pointed out that small sampling size left their study underpowered for testing the learning effect. Unsurprisingly, all of these four papers were among the experimental studies. However, none of the other six papers out of ten experimental studies in the corpus reported either calculating or considering a statistical power analysis while planning their research.

Transfer of learning was assessed only in four papers in the corpus, all of which are experimental studies (Chase & Abrahamson, 2015; Malinverni et al., 2012; Moher, 2009; Ruan et al., 2020). In Malinverni et al. (2012), bodily experience became the tool for construction of meaning, by allowing the transfer from concrete experience (gravity experienced in a slide) to explicit knowledge (Archimedes principle), with reference to Broaders, Cook, Mitchell, and Goldin-Meadow (2007). Ruan et al. (2020) measured delayed retention (the extent to which essential information is remembered) with a follow-up activity after a month and transfer of ability to apply gained knowledge and skills to solve new problems, with reference to an evidence-based approach to games for learning (Mayer, 2014). Similarly, in Moher (2009), the learners' retention and ability to transfer their understandings of a concept to a new domain was assessed with reference to Chen and Klahr (1999), while Chase and Abrahamson (2015) measured for the application of learned skills (transfer), but with no reference. In addition to these, four papers briefly mentioned transfer of learning; two experimental studies (Johnson et al., 2016a; Tan et al., 2016), two design-oriented research which uses both qualitative and quantitative research approaches (Antle et al., 2011; Beheshti et al., 2019). Transfer of learning in these four papers has not been applied or assessed, but only referred to as a limitation or future direction of the study.

Given that there are different types of experimental studies (i.e. true experiment and quasi-experimental) in quantitative research (Creswell & Creswell, 2017), we also coded as to whether this distinction was explicit in the papers. Only two experimental papers (Chase & Abrahamson, 2015; Tan et al., 2016) in the corpus explicitly identified their study as being quasi-experimental by clearly explaining that true-experimental design was not possible in their study as randomization of participants to groups was impractical in a classroom setting. For the rest of the experimental papers we further coded as to whether the sampling or the recruitment of participants procedure was random or non-random. We found four experimental studies (Flores-Gallegos et al., 2021; Malinverni et al., 2012; Michaelis & Mutlu, 2019; Ruan et al., 2020) in which the participants were randomly assigned to conditions in between-subject study. However, none of these four papers explicitly used the term *true experiment* to identify their study in the paper. We also found some papers in which the randomization procedure was not clear enough. For instance, in one experimental study (Apostolellis & Bowman, 2015) the authors explained that participants in their study were more democratic about who would get the control and did not always follow the intended random assignment. We reckon that this study was also non-randomized and thus quasi-experimental, as being implicitly described. Furthermore, in another experimental study (Westlund et al., 2017), there was no control group but two different robot conditions were compared within-subject in a counterbalanced order, yet the randomization strategy was not explained clearly in the paper. The remaining two experimental studies (Apostolellis & Bowman, 2016a; Johnson et al., 2016b) have not mentioned randomization at all, and thus again it was not possible to code these papers as to whether they were true or quasi-experimental studies.

4.2.2. Qualitative research approach

We found 12 studies in our corpus which adopted a solely qualitative research approach to assess learning (Beheshti et al., 2019; Bell & Davis, 2016; Garzotto & Bordogna, 2010; Ho et al., 2021; Keifert et al., 2017; Lamberty et al., 2011; Rick et al., 2011; Tseng et al., 2021; Vartiainen et al., 2020, 2021; Wan et al., 2020; Zimmermann-Niefield et al., 2020). Given the experimental studies also partly used qualitative methods, this number indicates that qualitative methods are the most commonly used approach to assess learning in CCI.

According to Creswell and Creswell (2017), types of qualitative research are e.g. narrative, phenomenology, ethnography, case study, grounded theory, participatory research. In our corpus, eight of the studies clearly characterized themselves as being a case study (Beheshti et al., 2019; Garzotto & Bordogna, 2010; Keifert et al., 2017; Rick et al., 2011; Tseng et al., 2021; Vartiainen et al., 2020, 2021; Wan et al., 2020) earlier than the Methods section except for one (Garzotto & Bordogna, 2010) which identified the study as a case study in the Discussion section. One of these case studies (Vartiainen et al., 2020) was also described as taking a participatory research approach to participatory learning. One of the qualitative studies was categorized as narrative summary using a grounded theory approach (Bell & Davis, 2016) to assess learning. Three qualitative studies did not mention their approach to assessing learning, thus it was not possible to categorize the characteristics of their research design (Ho et al., 2021; Lamberty et al., 2011; Zimmermann-Niefield et al., 2020). As mentioned above, there are also four studies that used a qualitative approach along with quantitative methods in their research to measure learning (Antle et al., 2011; Bartoli et al., 2013; Guo et al., 2016; Moher, 2009).

Even though reporting the recruitment strategy is as important in qualitative research, only one qualitative research paper in our corpus (Ho et al., 2021) mentioned about the randomization strategy in their recruitment procedure. Surprisingly, two of the qualitative studies introduced a hypothesis in the paper (Lamberty et al., 2011; Wan et al., 2020), but later in the Discussion section clarified that the hypothesis was not tested in the study.

4.2.3. Mixed-methods research approach

We found three papers in our corpus (Lee-Cultura et al., 2021; Pantic et al., 2016; Shimoda et al., 2013) which adopted a mixed-methods approach for assessing learning. According to Creswell and Creswell (2017) there are three types of mixed-methods research design, namely the convergent design, the explanatory sequential design, and the exploratory sequential design. We analyzed the three papers in terms of these three types. Lee-Cultura et al. described their analytical approach as being exploratory factor and convergent validity. The researchers analyzed children's play and problem solving behaviors based on multi-modal data (i.e., physiological, skeletal) combined with task performance (i.e. number of correct/incorrect matches) to understand the development in their geometry skills.

In Pantic et al. (2016), the researchers adopted an exploratory mixed methods approach to capture one girl's progress in programming skills in a Scratch summer camp. They carried out analysis across multiple forms of data, from traditional observational and artifact documentation to frequent, computationally generated saved data in several categories of programming concepts. They also analyzed changes, such as the number of programming blocks or art changes over time and her project outcomes by looking at four aspects: (i) project functionality, (ii) requirement satisfaction, (iii) changes introduced from one save to the other, and (iv) identification of what she spent her time doing. They noted which concepts the girl struggled with and when she made breakthroughs in learning. They combined transcripts and observational data in several iterative cycles to present the contexts in which key learning moments occurred, what led up to and followed those moments, and a fuller picture of her learning experiences.

The third, and final, paper that adopted a mixed methods approach in our corpus was (Shimoda et al., 2013), in which the researchers combine quantitative and qualitative methods to assess learning along with design-based research to evaluate the system that they are developing as a learning tool by looking at children's learning. A mixed-methods approach was used to assess three critical aspects of children's scientific inquiry learning: the inquiry cycle, cognitive and social factors, and metacognitive factors, science understanding, reasoning carefully, and inventiveness. By doing so, the researchers measured and assessed the development of students' understanding of and ability to apply the inquiry science concepts in collaborative learning opportunities. However, the researchers identified the overall methodology in their research as design-based research with particular reference to Brown (1992a) and Collins et al. (2004). In this paper, the researchers also noted that their evaluation of this system as a learning tool was not intended to be a controlled, empirical study, but rather a feasibility study of the usefulness of the Web of Inquiry in natural classroom environments that include a mix of teachers, students, curricula, and implementations.

None of the mixed methods studies above introduced and tested a hypothesis in their paper, nor mentioned a randomization strategy in the recruitment procedure.

4.2.4. Design-based research approach

The overall impression from the analyzed papers is that design-based research (Collins et al., 2004) where an inventive technological solution targeting learning is developed and tested is very common in CCI. However, when looking closer, only four of the papers in the corpus (Bell & Davis, 2016; Chase & Abrahamson, 2015; Shimoda et al., 2013; Vartiainen et al., 2021), make an explicit reference to commonly cited literature about educational design research. Further, several papers, e.g., Bonsignore et al. (2013), Gourlet and Decortis (2018), Katterfeldt, Cukurova, Spikol, and Cuartielles (2018), Yip et al. (2014) and Wyeth and MacColl

(2010) that were excluded due to not presenting any assessment of learning mention design-based research.

The remaining 26 papers were analyzed looking at if they present work that can be viewed as being design-based research even if it is not mentioned in the papers. Papers that were about the design, development and testing of a technological learning invention, where the context was a real-life learning environment were classified as cases of design-based research. Papers that were excluded were, for instance, papers that evaluated an existing technology without doing any design (e.g., Bartoli et al., 2013), and papers that did design a learning intervention but where the evaluation was only carried out in a lab context (e,g., Lee-Cultura et al., 2021; Ruan et al., 2020). The analysis showed that 11 papers did something that can be described as design research without making any reference to the concept. The papers were (Antle et al., 2011; Apostolellis & Bowman, 2015, 2016a; Garzotto & Bordogna, 2010; Guo et al., 2016; Johnson et al., 2016a; Keifert et al., 2017; Lamberty et al., 2011; Moher, 2009; Rick et al., 2011; Tan et al., 2016). Thus, in total half of the papers either explicitly or implicitly do (educational) design-based research employing learning artifact-centered evaluations that use artifacts in the experimental process. However, we found only three studies in which the DBR approach was utilized partly to assess a learning gain through observational techniques to investigate; how students externalize their prior knowledge and experiences through writing or artwork in Vartiainen et al. (2021), the ease in learning while using different types of interfaces in Antle et al. (2011), teens' conceptions of their roles and progress in the science program through PD activities (Bell & Davis, 2016).

4.3. Research design in assessing learning

Pre- and post-test for assessing learning were rather common: 22 studies conducted both pre- and post measures; 15 in the form of a skill test (Antle et al., 2011; Apostolellis & Bowman, 2015, 2016a; Bartoli et al., 2013; Flores-Gallegos et al., 2021; Lamberty et al., 2011; Malinverni et al., 2012; Moher, 2009; Rick et al., 2011; Ruan et al., 2020; Shimoda et al., 2013; Tan et al., 2016; Tseng et al., 2021; Vartiainen et al., 2021; Wan et al., 2020; Westlund et al., 2017), two in the form of interviews (Guo et al., 2016; Ho et al., 2021), and one in a self-report measure (Tisza & Markopoulos, 2021). Several studies conducted only a post-measure, one of which administered only a post-interview (Johnson et al., 2016a), one only a post-test to see a transfer (Chase & Abrahamson, 2015), and one conducted only post-observation (Beheshti et al., 2019).

The use of a control group was less common, 9 studies conducted a controlled study in which children received one of the conditions (Apostolellis & Bowman, 2015, 2016a; Beheshti et al., 2019; Chase & Abrahamson, 2015; Flores-Gallegos et al., 2021; Johnson et al., 2016a; Malinverni et al., 2012; Michaelis & Mutlu, 2019; Tan et al., 2016), and in two studies children received both conditions in a counterbalanced order (Ruan et al., 2020; Westlund et al., 2017) in a comparative approach.

In regards to criteria for learning (Schunk, 2012), a combination of all three criteria (administering pre-post tests, controlled group, and examining transfer of learning) were only found in three papers though; (Apostolellis & Bowman, 2016a; Malinverni et al., 2012; Ruan et al., 2020). These studies both have controlled vs. experimental conditions which meet the third criteria of learning. While the dependent variable of learning in Apostolellis and Bowman (2016a) is retention of content based on pre-test and delayed post-test which meets the first and second criteria of learning, Malinverni et al. (2012) compares the learning in two different contexts, thereby considering transfer of learning as the third criteria of learning. Similarly Ruan et al. conducted pre-post task and a pre-study survey that included a math anxiety scale survey and performed a between-subjects lab study to evaluate the three variants of the learning platform compared to the control and assessed the transfer of learning (Ruan et al., 2020).

Derived from the 11 definitions (ranging from One-Shot Case Study to A-B-A Single Subject Design) found in Creswell and Creswell pp. 231–41. the results for the characteristics of research design in the papers were as follows: The studies that adopted a qualitative approach to assess learning utilized either One-Shot Case Study (found in 5 papers; Bell & Davis, 2016; Garzotto & Bordogna, 2010; Keifert et al., 2017; Vartiainen et al., 2020; Zimmermann-Niefield et al., 2020) or One-Group Pretest-Posttest Design (found in 7 papers; Ho et al., 2021; Lamberty et al., 2011; Moher, 2009; Rick et al., 2011; Tseng et al., 2021; Vartiainen et al., 2021; Wan et al., 2020). There were more variations in terms of research design in quantitative studies in which Pretest-Posttest Control-Group Design was most commonly used (found in six out of 10 experimental studies Apostolellis & Bowman, 2015, 2016a; Flores-Gallegos et al., 2021; Michaelis & Mutlu, 2019; Ruan et al., 2020; Tan et al., 2016), two studies used One-Group Pretest-Posttest Design (Tisza & Markopoulos, 2021; Westlund et al., 2017), one experimental study used Alternative Treatment Posttest-Only With Nonequivalent Groups Design (in Johnson et al., 2016a), and one study used Solomon Four-Group Design (see Malinverni et al., 2012). The studies that used both qualitative and quantitative methods in their study used One-Group Pretest-Posttest Design (in Antle et al., 2011; Guo et al., 2016), Posttest-Only Control-Group Design (in Beheshti et al., 2019), and Single-Group Interrupted Time-Series Design (see Bartoli et al., 2013 in which children's attention variable was measured before, during and after the treatment). When we look at the research procedure design in the mixed methods papers, one of them used One-Group Pretest-Posttest Design (in Shimoda et al., 2013) and two papers used A-B-A Single-Subject Design (see Lee-Cultura et al., 2021; Pantic et al., 2016 in which neither pre-post data nor a controlled condition was administered, but complementary dependent variables were collected as source of data from individual participants to inform a specific aspect of learning). It is important to note that, the research design was only explicitly described in Malinverni et al. (2012) which used Solomon Four-Group Design.

4.4. Types of methods used to assess learning

For an overview of types of methods used to assess learning, see Table 3. In this table, only the methods pertaining to the assessment of learning used in the papers were included, while the methods for/of evaluation of the system experience or design were excluded.

4.4.1. Types of assessment methods in qualitative approach

Direct observations through video analysis. This type of method were employed to assess learning such as by capturing children's facial expressions and guided participation in machine learning activities (Vartiainen et al., 2020); investigating disabled children's understanding of temporal concepts in Garzotto and Bordogna (2010); conducting field trials to investigate the children's interaction with each other and creations while solving fraction-based problems (Lamberty et al., 2011); or observing a change in behavior (Keifert et al., 2017).

Direct observations — design oriented research. These studies combine multiple data sources with video recordings to understand children's learning such as emerging data-driven reasoning and how their evolving conceptual knowledge was contextualized through the co-design process and children's own app designs in Vartiainen et al. (2021); identifying learning outcomes associated with a series of participatory design activities in Bell and Davis (2016); children's brainstorming sessions in design workshops (Tseng et al., 2021); video recordings of dyad interactions with each other and the interface, along with researcher field notes and device log files that capture all interactions with the gadget to capture the parent and child body language, facial expressions, and gestures to compare content learning occurred in haptic and non-haptic versions in Beheshti et al. (2019).

Written responses — worksheets during workshops. Collecting children's brainstorming worksheets (Tseng et al., 2021).

Oral responses — follow-up questions. Supporting students to externalize their prior knowledge and evolving ideas on technology along with some follow-up questions about machine-learning concepts based on the intervention in Vartiainen et al. (2021).

Ratings by others — interviews with adults. The interviews conducted with educators, children's parents, and therapists, to evaluate the actual benefits gained by children and identify the issues that need to be addressed in Garzotto and Bordogna (2010), parents' views on a specific skill (e.g. counting and strategies such as pointing and finger counting in Ho et al., 2021).

Self-reports — questionnaires with open-ended questions for children. The methods take forms of e.g. pre- and post-workshop surveys were used to understand how participants' knowledge of machine learning changed as a result of the workshop in Tseng et al. (2021); pre-post questionnaires with written answers to capture children's understanding of machine learning concepts (e.g. clustering, similarity comparison, center point, k means clustering process and choosing an appropriate k number, and sense-making of patterns) in Wan et al. (2020).

Self-reports — questionnaires with adults. The questionnaires carried out to capture parents' views on children's specific skills (e.g. counting and strategies such as pointing and finger counting occurred in pre- and post activity in Ho et al., 2021).

Self-reports — interviews with children. Children were interviewed about their e.g., understandings of a mechanism and (task-specific) strategy arising from the preceding session in Moher, Liberati, Tetzlaff, and Altman (2009), or a concept such as evolution in Guo et al. (2016); semi-structured interviews were also conducted to investigate how children formed theories of machine learning in Zimmermann-Niefield et al. (2020).

Self-reports — interviews with parents or experts. Semi-structured interviews and three focus groups involving regular teachers, educators specialized in children with special needs, the school principal, and a language pathologist that assist teachers and children at school (Garzotto & Bordogna, 2010); interviews with families on their perceptions of their children's math learning through a game (Ho et al., 2021).

4.4.2. Types of assessment methods in quantitative approach

Direct observations — quantified behaviors. In quantitative studies observational notes were used to e.g., compare conditions to assess children's understanding of science principles through bodily experience in Malinverni et al. (2012), fun factor linked to playful learning in Tisza and Markopoulos (2021), or to analyze non-verbal cues such as children's gestural input (Beheshti et al., 2019; Malinverni et al., 2012; Zimmermann-Niefield et al., 2020).

Written responses — level of knowledge using an existing instrument or scale. Existing inventories are one of the most common tools used to measure learning in quantitative approach e.g. Social Skills Knowledge Test in Tan et al. (2016); in Matute, Rosselli, Ardila, and Ostrosky-Solís (2007) the Wechsler Intelligence Scale for Children (WISC-IV) standardized to Mexico and the Children Neuropsychological Scale (ENI-2) were used to specifically assess the areas of reading, writing, arithmetic, and visual attention; in Flores-Gallegos et al. (2021) they use TOVA (Test Of Variables of Attention) (Greenberg, 1991) a continuous performance test in

Table 3
Methods from the corpus used for assessing learning, divided into research approaches: qualitative (QL), quantitative (QT), design-based (DBR) and mixed-method (MM) approach.

Method	Research approach
Direct observations	QL: video recordings (Garzotto & Bordogna, 2010; Ho et al., 2021; Keifert et al., 2017; Lamberty et al., 2011; Moher, 2009; Vartiainen et al., 2020; Zimmermann-Niefield et al., 2020); DBR: Antle et al. (2011), Bell and Davis (2016) and Vartiainen et al. (2021) QT: behaviors or (non)verbal cues (Apostolellis & Bowman, 2015; Bartoli et al., 2013; Malinverni & Burguès, 2015; Westlund et al., 2017) MM: with Written Responses (Shimoda et al., 2013)
Written responses	QL: brainstorming worksheets (Tseng et al., 2021) QT: existing instruments or quizes (Bartoli et al., 2013; Flores-Gallegos et al., 2021; Malinverni et al., 2012; Michaelis & Mutlu, 2019; Ruan et al., 2020; Tan et al., 2016)
Oral responses	QL: follow-up questions about a concept (Antle et al., 2011; Guo et al., 2016; Vartiainen et al., 2021) QT: articulated understanding (Chase & Abrahamson, 2015)
Ratings by others	QL: parents' or therapists' views (Bartoli et al., 2013; Garzotto & Bordogna, 2010; Ho et al., 2021) QT: learning gains (Antle et al., 2011; Apostolellis & Bowman, 2015, 2016a; Chase & Abrahamson, 2015; Johnson et al., 2016a; Malinverni et al., 2012; Michaelis & Mutlu, 2019; Wan et al., 2020); task performance (Beheshti et al., 2019; Moher, 2009; Rick et al., 2011; Tisza & Markopoulos, 2021; Westlund et al., 2017); skills (Bartoli et al., 2013; Flores-Gallegos et al., 2021; Johnson et al., 2016b; Tan et al., 2016; Tisza & Markopoulos, 2021) MM: with Direct observations (Lee-Cultura et al., 2021)
Self-reports:Questionnaires	QL: open-ended questions children (Lamberty et al., 2011; Michaelis & Mutlu, 2019; Tseng et al., 2021; Wan et al., 2020); open-ended questions parents (Ho et al., 2021) QT: perceived knowledge (Apostolellis & Bowman, 2015; Guo et al., 2016; Michaelis & Mutlu, 2019; Tisza & Markopoulos, 2021)
Self-reports:Interviews	QL: with children (Guo et al., 2016; Lamberty et al., 2011; Moher, 2009; Vartiainen et al., 2020; Zimmermann-Niefield et al., 2020); with parents or educators (Garzotto & Bordogna, 2010; Ho et al., 2021; Shimoda et al., 2013)
Self-reports:Stimulated recalls	QT: perceived learning (Tisza & Markopoulos, 2021); retention of knowledge (Apostolellis & Bowman, 2016a; Moher, 2009; Ruan et al., 2020)
Self-reports: Think-alouds, Dialogues	QL: personal comments (Michaelis & Mutlu, 2019; Vartiainen et al., 2020) MM: with Direct Observations and artifact documentation (Pantic et al., 2016)

the visual,modality that evaluates selective visual attention and inhibition, Piers Harris self-concept scale (PH) (Fierro, Godoy, & Cardenal, 1987; Piers, Herzberg, & Harris, 1969) a scale centered on quantifying the self-perception of the participant based on areas of intellectual performance, sociability, anxiety, happiness, physical perception, and perception of one's own behavior, Biomechanical tasks (BM) consisted of four motor activities: standing on one foot with eyes open, standing on one foot with eyes closed, walking in a straight line with one foot in front of the other and lateral jumps with the feet together; standardized Bell test to measure attention skills in children with autism was used in Bartoli et al. (2013); math knowledge based on questions drawn from Khan Academy, and Singapore Math Books, with selection guided by the Common Core State Standards in Ruan et al. (2020); or the content knowledge tests consisted of multiple-choice and open ended questions being balanced in complexity from basic comprehension to synthesis and align with the science content in the treatment in Michaelis and Mutlu (2019).

Oral responses — participants' articulated understanding. Quantitive studies also utilized oral responses to translate into quantified data e.g., expressions and behaviors that indicated participants' development of subjective transparency of the algebra conceptual system (moments when the participants articulated understanding of the intervention) in Chase and Abrahamson (2015); students' metacognitive awareness of their learning such as appreciation of their community of practice in Bell and Davis (2016); tracing students' work to move from observations to expressed ideas and finally to consensus rules relating states of matter in Keifert et al. (2017).

Ratings by others — Scoring learning outcomes. The most common way to rate learning outcome is to measure the score difference between a pre-test and a post-test. There are also other quantitative rating methods used such as using rubrics to translate qualitative data into quantitative data related to learning outcomes along with calculating learning gain, which is the difference between pre-test and post-test scores (Antle et al., 2011); compare learning under two conditions (Apostolellis & Bowman, 2016b; Chase & Abrahamson, 2015); knowledge acquisition about principles of computing measured by a multiple choice test in Johnson et al. (2016b) or pre- and post survey on AI knowledge in Wan et al. (2020); comparing robot conditions for content knowledge and situational interest in Michaelis and Mutlu (2019).

Ratings by others — Task performance. There are various methods for quantifying the task performance found in CCI e.g. measuring the use of non-verbal cues such as eye-gaze and word recall in Westlund et al. (2017); engagement in a learning behavior or process by looking at task completion and group dynamics in Rick et al. (2011); use of verbal and non-verbal cues while interacting with haptic and non-haptic application of an exhibit (Beheshti et al., 2019); children's attitude about coding by using a Smiley-face scale (Tisza & Markopoulos, 2021).

Self-reports — Questionnaires. Asking children to rate their knowledge about a concept by using Likert scale before and after the intervention in Moher et al. (2009);

Self-reports — Stimulated recalls. This takes a form of measuring perceived learning (e.g. adopted from Papavlasopoulou, Sharma, & Giannakos, 2018) and delight using age appropriate Likert-type scale (de Leeuw, 2011; Mellor & Moore, 2013) in Tisza and Markopoulos (2021), or retention of knowledge and children's ability to transfer their understandings to a new domain (Apostolellis & Bowman, 2016a; Moher, 2009; Ruan et al., 2020).

4.4.3. Types of assessment methods in mixed methods studies

Combining direct observations with written responses. Testing the development of students' understanding of and ability to apply the inquiry science concepts through pre- and post-tests within and between subjects factor in Shimoda et al. (2013).

Combining direct observations with ratings by others. Collective analysis of multi-modal data to evaluate play and problem solving behaviors while doing a motion-based sorting task targeting geometry skill development in Lee-Cultura et al. (2021) and scoring children's sorting task performance by looking at correct matches counting the number of times the child placed a card in the correct box in Lee-Cultura et al. (2021).

Combining direct observations with self-reports - Think-alouds and artifact documentation. Applying a microgenetic approach to analysis across multiple forms of data, from traditional observational and artifact documentation to frequent, computationally generated save data for analyzing trajectory of one child's learning in Scratch Camp in Pantic et al. (2016).

4.5. Learning outcomes targeted in assessment

In this section we present what behaviors or skills were addressed for change and how change was assessed with regards to learning. We found three main categories for assessment: (1) assessing the change in children's understanding of a concept, (2) development in a skill, or (3) task completion or performance.

Fifteen papers assessed the change in children's understanding of a concept or phenomena such as knowledge of sustainability (Antle et al., 2011), conceptual knowledge of the domain (i.e. olive oil production and collaborative tasks being administered) (Apostolellis & Bowman, 2016b), scientific concepts (i.e., Archimedes principle Malinverni et al., 2012, electricity through haptic controller Beheshti et al., 2019, states of a matter Keifert et al., 2017, evolution Guo et al., 2016 learning science content from different types of robots Michaelis & Mutlu, 2019, algebra concepts Chase & Abrahamson, 2015, scientific inquiry Shimoda et al., 2013, learning about the scientific principle of "control of variables" Moher, 2009), machine learning concepts or AI knowledge (Tseng et al., 2021; Vartiainen et al., 2020, 2021; Wan et al., 2020), or ability to explain mechanisms (i.e. how a maker kit as an input/output device work Johnson et al., 2016b).

Ten papers assessed the development of a desired skill such as social skills, knowledge acquisition of social skills in general (Tan et al., 2016), social presence (Apostolellis & Bowman, 2015), negotiation in collaborative tasks (Apostolellis & Bowman, 2016a), math skills (Ho et al., 2021; Lamberty et al., 2011), attention skills (Bartoli et al., 2013), communication skills (i.e., use of non-verbal cues and word recall Westlund et al., 2017, vocabulary comprehension in children with special skills Garzotto & Bordogna, 2010), perceived coding skills (Tisza & Markopoulos, 2021) through self evaluation, or identifying learning outcomes linked to PD activities (Bell & Davis, 2016).

Nine papers examined a task completion or performance as an indicator of learning such as reading performance by looking at accuracy, speed, and comprehension (Flores-Gallegos et al., 2021), problem-solving behavior in a motion-based sorting task (Lee-Cultura et al., 2021), on and off-task conversation during engagement (Moher, 2009), a math anxiety scale survey and math performance before and after the treatment (Ruan et al., 2020), parent–child dyads' performance on the prediction of a math component as a gauge for their learning (Beheshti et al., 2019), group dynamics as a complementary to task completion (Rick et al., 2011), trajectory of learning to code through a microgenetic analysis (Pantic et al., 2016), on-task behavior and duration for creating machine learning models (Zimmermann-Niefield et al., 2020), or knowledge acquisition measured by performance on a test is the standard way to provide a quantitative measure of what has been learned (Johnson et al., 2016b).

Hereby, it is important to note that in some papers there are multiple learning goals being addressed and measured. That is why some of the papers occur more than once in the categories above.

5. Discussion and future work

Learning gains from technology use can lead to investments in technologies, which is why an increased research focus on assessment of learning as part of evaluating technology in educational settings can help educators and politicians to make research based informed choices. However, in coding the papers for assessment of learning as the final criteria for inclusion, we found that many papers (24 out of 54) do not assess learning and were therefore excluded from the final corpus. Moreover, two-thirds of the included papers have been published in the last five years. Thus, the very limited number of papers that ended up in our corpus indicate that assessment of learning outcomes is just starting to scratch the surface in CCI. Below we address and discuss some aspects for improvement in CCI research to generate evidence on learning outcomes.

5.1. Clarification of research aim

Some authors of excluded papers explicitly state that they have not studied learning (Fan & Antle, 2020; Li et al., 2020; Tisza & Markopoulos, 2021; Yip et al., 2014), while other excluded papers mostly evaluate situations for learning, although this is not clear from the introduction of the study. Many of these studies manifest that the goal is to facilitate, support, scaffold learning, but the methodology rather shows mediation of a learning situation, which leads to unclear assessment criteria for learning (Eriksson et al., 2022). Our hope for the future is that, when including assessment of learning as part of the evaluation of some technology, we should be explicit about the assessment criteria for learning.

Generally, in the papers included in the corpus, the research aim is sometimes obscure, or intermingled with various study goals etc. A common pattern that made us struggle while analyzing the papers was that the learning was assessed as part of the system evaluation or interaction experience, and factors such as engagement were measured as an indicator of interaction or experience rather than learning, albeit not clearly described in the paper. Inspired by Barendregt, Bekker, Börjesson, Eriksson, and Torgersson (2016) and Eriksson, Nilsson, Barendregt, and Nørgård (2021), we suggest to explicitly communicate the research goal and the learning goal for the children separately.

5.2. Clarification of research design

A common issue in the analyzed corpus was the lack of a clear characterization of the research design being adopted to assess learning. For instance, nine of the 12 qualitative studies in our corpus were characterized as case studies, whereas in the other three papers the research design has not been explicitly described at all. We believe that the rich variation in qualitative research approaches (e.g., grounded theory, phenomenology, ethnography, etc.) could have been utilized or reflected more in regards to learning in CCI.

When it comes to the next most common research approach for assessing learning, most of the quantitative studies (9/10) used experimental research to assess learning with a clear hypothesis. However, only two of these studies clearly explained that the study was quasi-experimental. Given that the distinction between true experiment and quasi-experiment is an important indicator for showing causality in learning in experimental studies (American Psychological Association, 2022a; Creswell & Creswell, 2017; Keppel & Wickens, 2004), this distinction needs to be clarified in the studies. For instance, we found that four papers mentioned carrying out a randomization process for participant assignment, albeit without explicitly identifying the study as a true experiment. Thus, the provided information does not necessarily clarify a randomized sampling procedure in recruiting the participants in the study. The participants may have been randomly assigned to groups within a non-random sampling procedure. Thus, it was not possible for us to classify any of these as a true experiment due to this unclear methodological description in the papers.

Furthermore, as the results show, there are very few studies that conducted a mixed methods approach to assess learning in CCI. A mixed methods approach, however, offers a strong view in understanding how learning occurs in a particular domain by investigating multiple dimensions of a specific data source. Thus, we believe CCI researchers could consider adopting more mixed-methods approaches to assess learning combining the strengths of qualitative and quantitative data with respect to a specific purpose.

In terms of the methodological rigor and clarity in assessment of learning, it is important to note that in some studies the formulation of the learning objectives and research design was obscure. For instance, in one paper the claim was to investigate children's understanding of a specific concept, however what has actually been assessed was children's programming approach to explain the phenomena which may also be regarded as task completion if the authors have not identified this as "understanding of a concept". Similarly, in another study, the performance on the prediction of a math component during the intervention may also be coded as understanding of a concept rather than performance if the authors had not claimed so. Furthermore, most of the time the basic characteristic of the research procedure was not described in the papers; especially the most commonly used ones such as One-Group Pretest-Posttest Design and Pretest–Posttest Control-Group Design. Even though we are aware of the difficulty in maintaining a standard language for all studies, we hope with this literature review, the terminology used in the analysis would shed a light for further research and help to leverage the methodological rigor, especially in terms of planning the research design to assess learning in our field.

In terms of planning the research design, if an analysis plan consists of detecting a significant association between variables of learning, a statistical power analysis can help to estimate a target sample size. However, we found only 3 papers out of 10 experimental studies that conducted power analysis in the research planning in the corpus. Moreover, even though many studies introduce their goal as being to facilitate, support, scaffold learning, their methodology rather demonstrate mediation of a learning situation, which leads to unclear assessment criteria for learning. However, a wide variety of legitimate scientific designs are available for educational research (National Academies Press, 2002), which means that experimental designs are very suitable for some studies, but are not indicated for others, and that experimental research can be productively complemented by other forms of enquiry (Taber, 2019).

The multidisciplinary nature of CCI research offers a fruitful field to further improve current theory and methods for evaluating technology mediating learning or adding learning as a parameter when evaluating technology in a learning context (Eriksson et al., 2022). Artificial Intelligence technologies for use in education are a good example of an emerging technology that will inevitably bring change, and we need reliable methodologies for evaluating their effectiveness in education (Cukurova & Luckin, 2018). Emerging technologies open up new horizons for the development of many new opportunities to understand, guide and enhance learning that were unimagined even a few years ago (Council, 2000), and as designers of such technologies we could certainly contribute. Also, we see that educational research is starting to adopt a design-based research approach both in designing and assessing educational interventions. However, in CCI we have used this design-based research approach for as long as the field has existed, and we therefore believe that CCI can contribute to theory and methods on design-based research methods when assessing learning as part of evaluating a technology.

5.3. Design-based research approach to understand learning as a phenomenon

In addition to the three main research approaches found in the literature, we also believe that the design-based research approach, which was found in half of the corpus, has a huge potential in understanding the complex and ungeneralizable nature of learning in CCI, especially in educational contexts (Brown, 1992a; Collins et al., 2004). However, we need more clear instances and reflections to realize how design-based research could be utilized for assessment of learning. Educational design-based research has many similarities with common approaches within CCI. It is typical both to do (technological) interventions, work together with schools in real life contexts, use an iterative approach to design and to try to develop guidelines that can be used by others. It is interesting, though, to see that despite the strong similarities, most papers in our corpus do not mention or reference educational design-based research. The impression is that the CCI field could gain from a raised awareness of the ideas behind DBR. In particular, DBR has a strong focus on knowledge development, and stresses that all interventions need to be thoroughly evaluated and their impact on all participants be assessed. Cobb, Confrey, diSessa, Lehrer, and Schauble describe how the knowledge development within DBR does not usually lead to grand theories of learning, but rather to knowledge constructs of an intermediate scope that are located somewhere between a specific system and a general theory (Cobb et al., 2003). This idea has a very direct connection to the concept of Intermediate knowledge used in HCI (see e.g Höök & Löwgren, 2012), which has also been suggested as a topic worth investigating further in CCI (Barendregt, Torgersson, Eriksson, & Börjesson, 2017).

5.4. Learning in/from extreme cases

When it comes to the learners being targeted, we found that five papers include some form of particular or extreme cases. One paper included children with reading disabilities (Flores-Gallegos et al., 2021), one with disabled children (Garzotto & Bordogna, 2010), one with autistic children (Bartoli et al., 2013), one partly included low socio-economic background (Apostolellis & Bowman, 2016b), and one paper on gifted children (Lamberty et al., 2011). This indicates that most of the papers in the corpus focus primarily on typically developing children. Inspired from studying extreme cases and designing for the very particular (Albert J. Mills, 2010; Bertelsen, Bødker, Eriksson, Hoggan, & Vermeulen, 2018; Stott & Hobden, 2015), such research approaches can help in informing how different types of learning using technology occur and can be supported. The extreme cases approach is employed when the purpose is to try to highlight the most unusual phenomena under investigation, rather than trying to tell something typical or average about the population in question (Albert J. Mills, 2010). This aligns with the view that learning is an individual activity: each of us learns at a different pace and has different cognitive abilities (Wing & Stanzione, 2016). Addressing the particular challenges of particular people in particular situations or activities has obvious societal relevance and impact on the group of particular people we work with. We see examples from extreme cases in learning science where e.g., a gifted high achiever was selected on the assumption that drawing attention to the characteristics of a successful learner may improve learning effectiveness of less successful learners (Stott & Hobden, 2015). In our corpus, we see something similar in a case study of one single girl's learning pathway (Pantic et al., 2016). However, in doing so, we need to balance between the value of the particular finding and the level of generalization, and accept that it may not always lead to results that can be applied outside the particular case nor have a large economic impact (Bertelsen et al., 2018).

5.5. Understanding multiple domains

The findings also show that the cognitive learning domain is dominant for assessing learning in the corpus. Further, we see that assessment of children's understanding of concepts related to STEM is most common, while complex phenomena such as arts, collaboration, equity, social justice, creativity, environmental matters, etc are less common or not present in the corpus. We see possibilities for future research to extend to other learning domains than the cognitive, as well to assess learning of complex concepts in extension to STEM.

Interactive technologies are increasingly playing a role in children's learning (Hourcade, 2022), both in traditional learning domains, but also in newer and more specific domains. In the review, we see that science and language learning as traditional learning domains hold a strong position, but that machine learning and programming as new learning domains are on the rise. In line with this, we also see an adoption of newer technologies in the corpus, such as robots, machine learning, and extended reality as complement to computers, large screens and mobile devices. We also see digital games as a popular means to support learning, building on early work by e.g. Fisch who outlined recommendations for the design of educational games (Fisch, 2005). Interactive technologies will most likely continue to play an increasing role in children's learning, and continue to be one of the major foci of research in CCI (Hourcade, 2022), both in traditional, newer and more specific learning domains and technologies. As designers of such technologies, there is an opportunity for CCI research to contribute with methods and theory on learning aspects of such technologies, if we chose to include assessment of learning in the evaluation of the technology.

5.6. Widening the scope and expanding the tendencies

The papers included in the corpus span the years 2009–2021, however 20/30 papers are published 2016 or later, which is why we might see a growing trend of assessing learning as part of the research within the field. Studies that assess understanding of a concept or phenomena are more in number (15/30) and they are more recent (half of them were conducted within the past 5 years and 5 of them within the past 3 years). Also these studies are more likely be formulated around formal learning (8/15) when compared to studies that assess development of a skill or on-task performance/behavior which were more carried out in informal learning. Furthermore, it seems as the year 2016 was also another turning point in terms of domains of learning; studies dealt with more with measuring social and communication skills (e.g. group dynamics, negotiation, collaboration, attention) whereas the focus was shifted to STEM and machine learning after 2016. Also, we have not encountered any examples that assess understanding of social skills, but social skills are more targeted to be developed or performed. Thus, changing perspectives may offer new opportunities to understand different domains of learning. Speaking of different domains of learning, we also have not encountered any studies that assess understanding of emotions, arts or music, or development of physical activity skills or performances. Thus, it would be interesting to widen the scope of learning domains in further studies.

5.7. Call for action: Text book

Screening the CCI text book literature on assessment, we typically find examples of evaluating technology with children, e.g. Druin (2009), Hourcade (2022) and Markopoulos et al. (2008), but as Cukurova and Luckin pointed out assessing learning as part of evaluating the technology is required and yet challenging to provide evidence for an impact. A few examples are mentioned but not elaborated in Hourcade, such as e.g the use of control group to show a statistically significant increase in standardized mathematics scores when using a web- and mobile-based mathematics game (McCarthy, Li, Tiu, & Atienza, 2013). Also, a book on experimental studies in learning technology and CCI was recently published (Giannakos, 2022). However, the promise of learning gains from computer use often entices educators and politicians to make investments in technologies (Hourcade, 2022), which is why it might be time for CCI research to increase such assessment in order to support investors in technology for learning in making research-based, informed choices. For this, we suggest a text book for CCI researchers on evaluating technology in learning situations and contexts, which would include methods and case studies for assessing learning as one way to evaluate technology.

5.8. Future avenues for including assessment of learning in CCI

- Clearer description of the terminology and basic characteristic of research design to increase the methodological rigor.
- Be explicit about the assessment criteria and dependent variable.
- Separate and communicate the research goal and the learning goal for the children explicitly to clarify the assessment criteria.
- Clarify causality between the effect of technology interaction on children's learning through true experiments.
- Conducting more mixed-methods to assess learning.
- Contribute to educational research theory and methods by including emerging technologies in research studies, and by applying a design-based research approach.
- Doing research on extreme cases and the very particular to help informing us on how different types of learning involving technology occurs and how it can be supported and mediated by technology.
- Assessment of children's understanding of concepts other than STEM, including complex phenomena such as arts, collaboration, equity, social justice, creativity, environmental matters, etc.
- A textbook on assessment of learning as part of evaluating technologies in educational settings for CCI research.

5.9. Limitations

This review has a number of limitations, particularly stemming from the title-abstract-keyword search strategy, the use of the stem learn*, and the two venues, which may have resulted in missing studies relevant to our aims. IDC and IJCCI are considered flagship venues within CCI (Eriksson et al., 2022; Giannakos, Papamitsiou et al., 2020; Van Mechelen, Baykal, Dindler, Eriksson, & Iversen, 2020), but are only representative of a selection of all published CCI research. Secondly, the corpus only contains publications that explicitly use the term learn*, well aware that many other terms are used such as e.g. "facilitate", "develop", "scaffold", "education", or "knowledge construction". Various other relevant terms may apply to assessment of learning, where some may more relate to, or be associated with, certain learning domains or research paradigms than others, and may therefore be disproportionately (under)represented in data. We are, however, specifically interested in the research approaches, methods and techniques to assess learning in CCI literature, and hence confined the search to capture studies that frame the focus specifically on learning phenomena. While, acknowledging that the inclusion of venues and search query introduces a selection bias, and that we might even skew the image of how learning has been assessed, we still think that this review presents clear insights on the assessment of learning in CCI research. Hopefully, these insights can inspire and guide CCI researchers publishing in a variety of venues.

6. Conclusion

This paper presents a semi-structured literature review focused on assessment of learning in the field of child–computer interaction (CCI) research. The review is based on a selection of 30 publications deriving from the CCI flagship venues the Interaction Design and Children (IDC) conference and the International Journal of Child–Computer Interaction (IJCCI). The results indicate that assessment of learning regarding transfer of learning and, controlled groups is rare, that the case studies using a qualitative research approach dominate the field whereas the mixed-methods approach is more rare. Furthermore, the findings showed that the basic characteristic of research design is rarely clearly defined which affects the methodological rigor and understanding causality of technology interaction in children's learning. Although the corpus is limited and can only be considered a sample, the paper maps out state of art of how to assess learning when evaluating some technology in CCI research, identifies gaps, and finally, we outline a number of future paths for the field.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

We warmly thank our colleagues and the anonymous reviewers for valuable feedback.

Selection and participation

No children participated in this work.

References

Ackermann, E. (1991). The " agency " model of transactions : Toward an understanding of children ' s theory of control.

Albert J. Mills, E. W. (Ed.), (2010). *Extreme cases*. SAGE Publications, Inc., http://dx.doi.org/10.4135/9781412957397.n142.

American Psychological Association (2022a). APA dictionary of psychology. URL: https://dictionary.apa.org/.

American Psychological Association (2022b). APA Dictionary of Psychology - dependent variable (DV). URL: https://dictionary.apa.org/dependent-variable.

Antle, A. N., Wise, A. F., & Nielsen, K. (2011). Towards utopia: Designing tangibles for learning. In *Proceedings of the 10th international conference on interaction design and children* IDC '11, (pp. 11–20). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/1999030.1999032.

Apostolellis, P., & Bowman, D. A. (2015). Small group learning with games in museums: Effects of interactivity as mediated by cultural differences. In *Proceedings of the 14th international conference on interaction design and children* IDC '15, (pp. 160–169). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/2771839.2771856.

Apostolellis, P., & Bowman, D. A. (2016a). Audience involvement and agency in digital games: Effects on learning, game experience, and social presence. In *Proceedings of the the 15th international conference on interaction design and children* IDC '16, (pp. 299–310). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/2930674.2930700.

Apostolellis, P., & Bowman, D. A. (2016b). Audience involvement and agency in digital games: Effects on learning, game experience, and social presence. In *Proceedings of the the 15th international conference on interaction design and children* IDC '16, (pp. 299–310). New York, NY, USA: ACM, http://dx.doi.org/10.1145/2930674.2930700, URL: http://doi.acm.org/10.1145/2930674.2930700, event-place: Manchester, United Kingdom.

Barendregt, W., Bekker, T. M., Börjesson, P., Eriksson, E., & Torgersson, O. (2016). Legitimate participation in the classroom context: Adding learning goals to participatory design. In *Proceedings of the the 15th international conference on interaction design and children* IDC '16, (pp. 167–174). New York, NY, USA: ACM, http://dx.doi.org/10.1145/2930674.2930686, URL: http://doi.acm.org/10.1145/2930674.2930686, event-place: Manchester, United Kingdom.

Barendregt, W., Torgersson, O., Eriksson, E., & Börjesson, P. (2017). Intermediate-level knowledge in child-computer interaction: A call for action. In *Proceedings of the 2017 conference on interaction design and children* IDC '17, (pp. 7–16). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/3078072.3079719.

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn?: A taxonomy for far transfer. *128*, 612–637. http://dx.doi.org/10.1037/0033-2909.128.4.612,

Bartoli, L., Corradi, C., Garzotto, F., & Valoriani, M. (2013). Exploring motion-based touchless games for autistic children's learning. In *Proceedings of the 12th international conference on interaction design and children* IDC '13, (pp. 102–111). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/2485760.2485774.

Baykal, G., Alaca, I. V., Yantaç, A., & Göksun, T. (2018). A review on complementary natures of tangible user interfaces (TUIs) and early spatial learning. *International Journal of Child-Computer Interaction*, *16*, 104–113. http://dx.doi.org/10.1016/j.ijcci.2018.01.003.

Becker, K. (2010). The clark-kozma debate in the 21-st century.

Beheshti, E., Borgos-Rodriguez, K., & Piper, A. M. (2019). Supporting parent-child collaborative learning through haptic feedback displays. In *Proceedings of the 18th ACM international conference on interaction design and children* IDC '19, (pp. 58–70). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/3311927.3323137.

Bekker, T., & Antle, A. N. (2011). *Developmentally situated design (DSD)*. ACM, http://dx.doi.org/10.1145/1978942.1979312.

Bell, A., & Davis, K. (2016). Learning through participatory design: Designing digital badges for and with teens. In *Proceedings of the the 15th international conference on interaction design and children* IDC '16, (pp. 218–229). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/2930674.2930705.

Bertelsen, O. W., Bødker, S., Eriksson, E., Hoggan, E., & Vermeulen, J. (2018). Beyond generalization: Research for the very particular. *Interactions*, *26*(1), 34–38. http://dx.doi.org/10.1145/3289425.

Bonsignore, E., Hansen, D., Kraus, K., Visconti, A., Ahn, J., & Druin, A. (2013). Playing for real: Designing alternate reality games for teenagers in learning contexts. In *Proceedings of the 12th international conference on interaction design and children* IDC '13, (pp. 237–246). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/2485760.2485788.

Boudon, R. (1969). Acta sociologica. *12*, 161–163. URL: http://www.jstor.org/stable/4193712,

Broaders, S. C., Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2007). Making children gesture brings out implicit knowledge and leads to learning. *Journal of Experimental Psychology: General*, *136*(4), 539–550.

Brown, A. (1992a). Design experiments : Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences*, *2*(2), 141–178.

Brown, A. L. (1992b). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *Journal of the Learning Sciences*, *2*(2), 141–178. http://dx.doi.org/10.1207/s15327809jls0202_2.

Bunting, L., af Segerstad, Y. H., & Barendregt, W. (2021). Swedish teachers' views on the use of personalised learning technologies for teaching children reading in the English classroom. *27*, 100236. http://dx.doi.org/10.1016/j.ijcci.2020.100236,

Chase, K., & Abrahamson, D. (2015). Reverse scaffolding: A constructivist design architecture for mathematics learning with educational technology. In *Proceedings of the 14th international conference on interaction design and children* IDC '15, (pp. 189–198). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/2771839.2771859.

Chen, Z., & Klahr, D. (1999). All other things being equal: Acquisition and transfer of the control of variables strategy. *Child Development*, *70*(5), 1098–1120.

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, *32*(1), 9–13. http://dx.doi.org/10.3102/0013189x032001009, URL: https://journals.sagepub.com/doi/abs/10.3102/0013189X032001009.

COE, C. o. E. Formal, non-formal and informal learning. URL: https://www.coe.int/en/web/lang-migrants/formal-non-formal-and-informal-learning,

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Routledge.

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *The Journal of the Learning Sciences*, *13*(1), 15–42.

Council, N. R. (2000). *How people learn: Brain, mind, experience & school*. National Academies Press, http://dx.doi.org/10.17226/9853.

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.

Cukurova, M., & Luckin, R. (2018). Measuring the impact of emerging technologies in education: A pragmatic approach. In *Springer international handbooks of education* (pp. 1–19). Springer International Publishing, http://dx.doi.org/10.1007/978-3-319-53803-7_81-1.

Cukurova, M., Luckin, R., & Clark-Wilson, A. (2019). Creating the golden triangle of evidence-informed education technology with EDUCATE. *British Journal of Educational Technology*, *50*(2), 490–504. http://dx.doi.org/10.1111/bjet.12727, URL: https://bera-journals.onlinelibrary.wiley.com/doi/abs/10.1111/bjet.12727. arXiv:https://bera-journals.onlinelibrary.wiley.com/doi/pdf/10.1111/bjet.12727.

de Leeuw, E. D. (2011). Improving data quality when surveying children and adolescents: Cognitive and social development and its.

Druin, A. (2009). *Mobile technology for children*.

Einarsson, Á. M., & Hertzum, M. (2020). How is learning scaffolded in library makerspaces? *International Journal of Child-Computer Interaction*, *26*, Article 100199. http://dx.doi.org/10.1016/j.ijcci.2020.100199.

Eriksson, E., Baykal, G. E., & Torgersson, O. (2022). The role of learning theory in child-computer interaction - A semi-systematic literature review. In *Interaction design and children* IDC '22, (pp. 50–68). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/3501712.3529728.

Eriksson, E., Nilsson, E. M., Barendregt, W., & Nørgård, R. T. (2021). Teaching values in design in higher education – towards a new normal.. In *Proceedings of conference on the ethical and social impacts of ICT – Ethicomp2021*. Ethicomp '21, Logrono, Spain: Universidad de la Rioja, Spain..

Eshach, H. (2007). Bridging in-school and out-of-school learning: Formal, non-formal, and informal education. *Journal of Science Education and Technology*, *16*(2), 171–190.

Fan, M., & Antle, A. N. (2020). An english language learning study with rural Chinese children using an augmented reality app. In *Proceedings of the interaction design and children conference* IDC '20, (pp. 385–397). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/3392063.3394409.

Feldman, R., & Garrison, M. (1993). *Understanding psychology*. New York: McGraw-Hill.

Fierro, A., Godoy, A., & Cardenal, V. (1987). Dimensiones, consistencia y estabilidad en la escala de autoconcepto "Piers-Harris". In *Póster presentado al II congreso de evaluación psicológica (Madrid*.

Fisch, S. M. (2005). Making educational computer games "educational". In *Proceedings of the 2005 conference on interaction design and children* IDC '05, (pp. 56–61). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/1109540.1109548.

Flores-Gallegos, R., Rodríguez-Leis, P., & Fernández, T. (2021). Effects of a virtual reality training program on visual attention and motor performance in children with reading learning disability. *International Journal of Child-Computer Interaction*, Article 100394. http://dx.doi.org/10.1016/j.ijcci.2021.100394.

Flynn, R. M., Kleinknecht, E., Ricker, A. A., & Blumberg, F. C. (2021). A narrative review of methods used to examine digital gaming impacts on learning and cognition during middle childhood. *International Journal of Child-Computer Interaction*, *30*, Article 100325. http://dx.doi.org/10.1016/j.ijcci.2021.100325.

Frauenberger, C., Good, J., Fitzpatrick, G., & Iversen, O. S. (2015). In pursuit of rigour and accountability in participatory design. *International Journal of Human-Computer Studies*, *74*, 93–106. http://dx.doi.org/10.1016/j.ijhcs.2014.09.004.

Gajda, A., Karwowski, M., & Beghetto, R. A. (2017). Creativity and academic achievement: A meta-analysis. *Journal of Educational Psychology*, *109*(2), 269.

Garzotto, F., & Bordogna, M. (2010). Paper-based multimedia interaction as learning tool for disabled children. In *Proceedings of the 9th international conference on interaction design and children* IDC '10, (pp. 79–88). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/1810543.1810553.

Giannakos, M. (2022). *Experimental studies in learning technology and child–computer interaction*. Springer International Publishing, http://dx.doi.org/10.1007/978-3-031-14350-2.

Giannakos, M. N., Horn, M. S., Read, J. C., & Markopoulos, P. (2020). Movement forward: The continued growth of Child–Computer Interaction research. *International Journal of Child-Computer Interaction*, *26*, Article 100204. http://dx.doi.org/10.1016/j.ijcci.2020.100204.

Giannakos, M., Papamitsiou, Z., Markopoulos, P., Read, J., & Hourcade, J. P. (2020). Mapping child–computer interaction research through co-word analysis. *International Journal of Child-Computer Interaction*, *23–24*, Article 100165. http://dx.doi.org/10.1016/j.ijcci.2020.100165, URL: https://www.sciencedirect.com/science/article/pii/S2212868920300015.

Gourlet, P., & Decortis, F. (2018). Prototyping a designerly learning through authentic making activities in elementary classrooms. *International Journal of Child-Computer Interaction*, *16*, 31–38. http://dx.doi.org/10.1016/j.ijcci.2017.11.002.

Greenberg, L. M. (1991). *T.O.V.A. interpretation manual*. Minneapolis, MN.

Guo, Y., Wagh, A., Brady, C., Levy, S. T., Horn, M. S., & Wilensky, U. (2016). Frogs to think with: Improving students' computational thinking and understanding of evolution in a code-first learning environment. In *Proceedings of the the 15th international conference on interaction design and children* IDC '16, (pp. 246–254). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/2930674.2930724.

Ho, H.-R., Cagiltay, B., White, N. T., Hubbard, E. M., & Mutlu, B. (2021). RoboMath: Designing a learning companion robot to support children's numerical skills. In *Interaction design and children* IDC '21, (pp. 283–293). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/3459990.3460709.

Höök, K., & Löwgren, J. (2012). Strong concepts: Intermediate-level knowledge in interaction design research. *ACM Transactions on Computer-Human Interaction*, *19*(3), Article 23. http://dx.doi.org/10.1145/2362364.2362371.

Hourcade, J. P. (2007). Interaction design and children. *Foundations and Trends® in Human-Computer Interaction*, *1*(4), 277–392. http://dx.doi.org/10.1561/1100000006.

Hourcade, J. P. (2022). *Child-computer interaction, second edition*. US.

Iversen, O. S., Smith, R. C., & Dindler, C. (2017). Child as protagonist: Expanding the role of children in participatory design. In *Proceedings of the 2017 conference on interaction design and children* IDC '17, (pp. 27–37). New York, NY, USA: ACM, http://dx.doi.org/10.1145/3078072.3079725, URL: http://doi.acm.org/10.1145/3078072.3079725, event-place: Stanford, California, USA.

Johnson, R., Shum, V., Rogers, Y., & Marquardt, N. (2016a). Make or shake: An empirical study of the value of making in learning about computing technology. In *Proceedings of the the 15th international conference on interaction design and children* IDC '16, (pp. 440–451). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/2930674.2930691.

Johnson, R., Shum, V., Rogers, Y., & Marquardt, N. (2016b). Make or shake: An empirical study of the value of making in learning about computing technology. In *Proceedings of the the 15th international conference on interaction design and children* IDC '16, (pp. 440–451). New York, NY, USA: ACM, http://dx.doi.org/10.1145/2930674.2930691, URL: http://doi.acm.org/10.1145/2930674.2930691, event-place: Manchester, United Kingdom.

Katterfeldt, E.-S., Cukurova, M., Spikol, D., & Cuartielles, D. (2018). Physical computing with plug-and-play toolkits:Key recommendations for collaborative learning implementations. *International Journal of Child-Computer Interaction*, *17*, 72–82. http://dx.doi.org/10.1016/j.ijcci.2018.03.002.

Keifert, D., Lee, C., Dahn, M., Illum, R., DeLiema, D., Enyedy, N., et al. (2017). Agency, embodiment, & affect during play in a mixed-reality learning environment. In *Proceedings of the 2017 conference on interaction design and children* IDC '17, (pp. 268–277). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/3078072.3079731.

Keppel, G., & Wickens, T. D. (2004). *Design and analysis: A researcher's handbook* (4th ed.). Upper Saddle River, NJ: Pearson.

Kolodner, J. L., Gray, J., & Fasse, B. B. (2003). Promoting transfer through case-based reasoning: Rituals and practices in learning by design classrooms. *Cognitive Science Quarterly*, *3*(2), 183–232.

Lamberty, K. K., Adams, S., Biatek, J., Froiland, K., & Lapham, J. (2011). Using a large display in the periphery to support children learning through design. In *Proceedings of the 10th international conference on interaction design and children* IDC '11, (pp. 62–71). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/1999030.1999038.

Lee-Cultura, S., Sharma, K., & Giannakos, M. (2021). Children's play and problem-solving in motion-based learning technologies using a multi-modal mixed methods approach. Article 100355. http://dx.doi.org/10.1016/j.ijcci.2021.100355.

Li, J., van der Spek, E. D., Yu, X., Hu, J., & Feijs, L. (2020). Exploring an augmented reality social learning game for elementary school students. In *Proceedings of the interaction design and children conference* IDC '20, (pp. 508–518). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/3392063.3394422.

Lyle, P., Korsgaard, H., & Bødker, S. (2020). What's in an ecology? A review of artifact, communicative, device and information ecologies. In *Proceedings of the 11th nordic conference on human-computer interaction: Shaping experiences, shaping society*. New York, NY, USA: Association for Computing Machinery, URL: https://doi.org/10.1145/3419249.3420185.

Malinverni, L., & Burguès, N. P. (2015). The medium matters: The impact of full-body interaction on the socio-affective aspects of collaboration. In *Proceedings of the 14th international conference on interaction design and children* IDC '15, (pp. 89–98). New York, NY, USA: ACM, http://dx.doi.org/10.1145/2771839.2771849, URL: http://doi.acm.org/10.1145/2771839.2771849, event-place: Boston, Massachusetts.

Malinverni, L., Silva, B. L., & Parés, N. (2012). Impact of embodied interaction on learning processes: Design and analysis of an educational application based on physical activity. In *Proceedings of the 11th international conference on interaction design and children* IDC '12, (pp. 60–69). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/2307096.2307104.

Manches, A., & Price, S. (2011). Designing learning representations around physical manipulation: Hands and objects. In *Proceedings of the 10th international conference on interaction design and children* IDC '11, (pp. 81–89). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/1999030.1999040.

Markopoulos, P., Read, J., MacFarlane, S., & Hoysniemi, J. (2008). *Evaluating children's interactive products*. Elsevier, http://dx.doi.org/10.1016/b978-0-12-374111-0.x0001-1.

Matute, E., Rosselli, M., Ardila, A., & Ostrosky-Solís, F. (2007). *Evaluación neuropsicológica infantil*. México: Manual Moderno.

Mayer, R. E. (2014). *Computer games for learning: An evidence-based approach*. MIT Press.

Mayers, A. (2013). *Introduction to statistics and SPSS in psychology*. London, England: Pearson Education.

McCarthy, B., Li, L., Tiu, M., & Atienza, S. (2013). PBS KIDS mathematics transmedia suites in preschool homes. In *Proceedings of the 12th international conference on interaction design and children* IDC '13, (pp. 128–136). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/2485760.2485777.

Mellor, D., & Moore, K. A. (2013). The use of likert scales with children. *Journal of Pediatric Psychology*, *39*(3), 369–379. http://dx.doi.org/10.1093/jpepsy/jst079, arXiv:https://academic.oup.com/jpepsy/article-pdf/39/3/369/3407451/jst079.pdf.

Michaelis, J. E., & Mutlu, B. (2019). Supporting interest in science learning with a social robot. In *Proceedings of the 18th ACM international conference on interaction design and children* IDC '19, (pp. 71–82). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/3311927.3323154.

Moher, T. (2009). Putting interference to work in the design of a whole-class learning activity. In *Proceedings of the 8th international conference on interaction design and children* IDC '09, (pp. 115–122). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/1551788.1551808.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *6*(7), Article e1000097. http://dx.doi.org/10.1371/journal.pmed.1000097.

(2002). *Scientific research in education*. National Academies Press, http://dx.doi.org/10.17226/10236.

OECD, O. f. E. C.-o., & Development Recognition of Non-formal and Informal Learning - Home. URL: https://www.oecd.org/fr/education/apprendre-au-dela-de-l-ecole/recognitionofnon-formalandinformallearning-home.htm,

Oranç, C., & Küntay, A. C. (2019). Learning from the real and the virtual worlds: Educational use of augmented reality in early childhood. *International Journal of Child-Computer Interaction*, *21*, 104–111. http://dx.doi.org/10.1016/j.ijcci.2019.06.002.

Pantic, K., Fields, D. A., & Quirke, L. (2016). Studying situated learning in a constructionist programming camp: A multimethod microgenetic analysis of one girl's learning pathway. In *Proceedings of the the 15th international conference on interaction design and children* IDC '16, (pp. 428–439). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/2930674.2930725.

Papavlasopoulou, S., Sharma, K., & Giannakos, M. N. (2018). How do you feel about learning to code? Investigating the effect of children's attitudes towards coding using eye-tracking. *International Journal of Child-Computer Interaction*, *17*, 50–60.

Papert, S. (1980). *Mindstorms : Children, computers, and powerful ideas* (p. 230). Basic Books, URL: https://dl.acm.org/citation.cfm?id=1095592.

Piers, E. V., Herzberg, D. S., & Harris, D. B. (1969). *The piers-harris children's self-concept scale*. Tennessee, USA: Western Psychological Services.

Plomp, T. (2013). Educational design research: An introduction. *Educational Design Research*, 11–50.

Popham, W. (2008). *Classroom assessment: What teachers need to know*. Boston: Allyn & Bacon.

Read, J. C., Horton, M., Sim, G., Gregory, P., Fitton, D., & Cassidy, B. (2013). CHECk. In *CHI '13 extended abstracts on human factors in computing systems on - CHI EA '13*. ACM Press, http://dx.doi.org/10.1145/2468356.2468391.

Rick, J., Marshall, P., & Yuill, N. (2011). Beyond one-size-fits-all: How interactive tabletops support collaborative learning. In *Proceedings of the 10th international conference on interaction design and children* IDC '11, (pp. 109–117). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/1999030.1999043.

Roskos, K., Brueck, J., & Lenhart, L. (2017). An analysis of e-book learning platforms: Affordances, architecture, functionality and analytics. *12*, 37–45. http://dx.doi.org/10.1016/j.ijcci.2017.01.003,

Ruan, S., He, J., Ying, R., Burkle, J., Hakim, D., Wang, A., et al. (2020). Supporting children's math learning with feedback-augmented narrative technology. In *Proceedings of the interaction design and children conference* IDC '20, (pp. 567–580). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/3392063.3394400.

Schunk, D. (2012). *Learning theories : An educational perspective*. Boston: Pearson.

Schwartz, D. L. (2005). Efficiency and innovation in transfer.

Shawn Green, C., Bavelier, D., Kramer, A. F., Vinogradov, S., Ansorge, U., Ball, K. K., et al. (2019). Improving methodological standards in behavioral interventions for cognitive enhancement. *Journal of Cognitive Enhancement*, *3*(1), 2–29.

Shimoda, T., White, B., Borge, M., & Frederiksen, J. (2013). Designing for science learning and collaborative discourse. In *Proceedings of the 12th international conference on interaction design and children* IDC '13, (pp. 247–256). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/2485760.2485782.

Shuell, T. J. (1986). Cognitive conceptions of learning. *Review of Educational Research*, *56*(4), 411–436. http://dx.doi.org/10.3102/00346543056004411.

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, *104*, 333–339. http://dx.doi.org/10.1016/j.jbusres.2019.07.039, URL: https://www.sciencedirect.com/science/article/pii/S0148296319304564.

Stott, A., & Hobden, P. A. (2015). Effective learning. *Gifted Child Quarterly*, *60*(1), 63–74. http://dx.doi.org/10.1177/0016986215611961.

Taber, K. S. (2019). Experimental research into teaching innovations: responding to methodological and ethical challenges. *Studies in Science Education*, *55*(1), 69–119. http://dx.doi.org/10.1080/03057267.2019.1658058.

Tan, J. L., Goh, D. H.-L., Ang, R. P., & Huan, V. S. (2016). Learning efficacy and user acceptance of a game-based social skills learning environment. *International Journal of Child-Computer Interaction*, *9–10*, 1–19. http://dx.doi.org/10.1016/j.ijcci.2016.09.001.

Tisza, G., & Markopoulos, P. (2021). Understanding the role of fun in learning to code. *International Journal of Child-Computer Interaction*, *28*, Article 100270. http://dx.doi.org/10.1016/j.ijcci.2021.100270.

Tisza, G., Papavlasopoulou, S., Christidou, D., Iivari, N., Kinnula, M., & Voulgari, I. (2020). Patterns in informal and non-formal science learning activities for children–A Europe-wide survey study. *International Journal of Child-Computer Interaction*, *25*, Article 100184. http://dx.doi.org/10.1016/j.ijcci.2020.100184.

Tisza, G., Papavlasopoulou, S., Christidou, D., Voulgari, I., Iivari, N., Giannakos, M. N., et al. (2019). The role of age and gender on implementing informal and non-formal science learning activities for children. In *FabLearn Europe '19*, *Proceedings of the fablearn Europe 2019 conference*. New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/3335055.3335065.

Troseth, G. L., & Strouse, G. A. (2017). Designing and using digital books for learning: The informative case of young children and video. *12*, 3–7. http://dx.doi.org/10.1016/j.ijcci.2016.12.002,

Tseng, T., Murai, Y., Freed, N., Gelosi, D., Ta, T. D., & Kawahara, Y. (2021). PlushPal: Storytelling with interactive plush toys and machine learning. In *Interaction design and children* IDC '21, (pp. 236–245). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/3459990.3460694.

Tuckman, B. (2012). *Conducting educational research*. Lanham, Md: Rowman & Littlefield Publishers.

Van Mechelen, M., Baykal, G. E., Dindler, C., Eriksson, E., & Iversen, O. S. (2020). 18 Years of ethics in child-computer interaction research: A systematic literature review. In *Proceedings of the interaction design and children conference* IDC '20, (pp. 161–183). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/3392063.3394407.

Vartiainen, H., Tedre, M., & Valtonen, T. (2020). Learning machine learning with very young children: Who is teaching whom? *International Journal of Child-Computer Interaction*, *25*, Article 100182. http://dx.doi.org/10.1016/j.ijcci.2020.100182.

Vartiainen, H., Toivonen, T., Jormanainen, I., Kahila, J., Tedre, M., & Valtonen, T. (2021). Machine learning for middle schoolers: Learning through data-driven design. *International Journal of Child-Computer Interaction*, *29*, Article 100281. http://dx.doi.org/10.1016/j.ijcci.2021.100281.

Wan, X., Zhou, X., Ye, Z., Mortensen, C. K., & Bai, Z. (2020). SmileyCluster: Supporting accessible machine learning in K-12 scientific discovery. In *Proceedings of the interaction design and children conference* IDC '20, (pp. 23–35). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/3392063.3394440.

Westlund, J. M. K., Dickens, L., Jeong, S., Harris, P. L., DeSteno, D., & Breazeal, C. L. (2017). Children use non-verbal cues to learn new words from robots as well as people. *International Journal of Child-Computer Interaction*, *13*, 1–9. http://dx.doi.org/10.1016/j.ijcci.2017.04.001.

Wing, J. M., & Stanzione, D. (2016). Progress in computational thinking, and expanding the HPC community. *Communications of the ACM*, *59*(7), 10–11. http://dx.doi.org/10.1145/2933410.

Wong, G., Greenhalgh, T., Westhorp, G., Buckingham, J., & Pawson, R. (2013). RAMESES publication standards: meta-narrative reviews. *11*. http://dx.doi.org/10.1186/1741-7015-11-20,

Wyeth, P., & MacColl, I. (2010). Noising around: Investigations in mobile learning. In *Proceedings of the 9th international conference on interaction design and children* IDC '10, (pp. 147–155). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/1810543.1810560.

Yip, J., Ahn, J., Clegg, T., Bonsignore, E., Pauw, D., & Gubbels, M. (2014). "It helped me do my science.": A case of designing social media technologies for children in science learning. In *Proceedings of the 2014 conference on interaction design and children* IDC '14, (pp. 155–164). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/2593968.2593969.

Zimmermann-Niefield, A., Polson, S., Moreno, C., & Shapiro, R. B. (2020). Youth making machine learning models for gesture-controlled interactive media. In *Proceedings of the interaction design and children conference* IDC '20, (pp. 63–74). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/3392063.3394438.

Update

International Journal of Child-Computer Interaction
Volume 41, Issue , September 2024, Page

DOI: https://doi.org/10.1016/j.ijcci.2024.100677

International Journal of Child-Computer Interaction 41 (2024) 100677

Contents lists available at ScienceDirect

International Journal of Child-Computer Interaction

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

Erratum regarding missing Declaration of Competing Interest statements in previously published articles

Declaration of Competing Interest statements were not included in the published version of the following articles that appeared in previous issues of International Journal of Child-Computer Interaction.

The appropriate Declaration/Competing Interest statements, provided by the Authors, are included below.

"How block-based, text-based, and hybrid block/text modalities shape novice programming practices" (International Journal of Child-Computer Interaction 17, (2018) 83–92) https://doi.org/10.1016/j.ijcci.2018.04.005. Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work in this paper.

"Socializers, achievers or both? Value-based roles of children in technology design projects" (International Journal of Child-Computer Interaction 17, (2018), 39–49) https://doi.org/10.1016/j.ijcci.2018.04.004. Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work in this paper.

"A comparative study into how pupils can play different roles in co-design activities" (International Journal of Child-Computer Interaction 17, (2018), 28–38) https://doi.org/10.1016/j.ijcci.2018.04.003 Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work in this paper.

"Deepening children's STEM learning through making and creative writing" (International Journal of Child-Computer Interaction 40, (2024), 100651) https://doi.org/10.1016/j.ijcci.2024.100651 Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work in this paper.

"Validating a performance assessment of computational thinking for early childhood using item response theory" (International Journal of Child-Computer Interaction 40, (2024), 100650) https://doi.org/10.1016/j.ijcci.2024.100650 Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work in this paper.

"The impact of near-peer virtual agents on computer science attitudes and collaborative dialogue" (International Journal of Child-Computer Interaction 40, (2024), 100646) https://doi.org/10.1016/j.ijcci.2024.100646 Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work in this paper.

"Children's approaches to solving puzzles in videogames" (International Journal of Child-Computer Interaction 40, (2024), 100635) https://doi.org/10.1016/j.ijcci.2024.100635 Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work in this paper.

"Making the invisible visible: Youth designs for teaching about technological and algorithmic bias" (International Journal of Child-Computer Interaction 39, (2024), 100634) https://doi.org/10.1016/j.ijcci.2024.100634 Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work in this paper.

"Children's engagement during shared reading of ebooks and paper books: A systematic review" (International Journal of Child-Computer Interaction 39, (2024), 100632) https://doi.org/10.1016/j.ijcci.2023.100632 Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work in this paper.

"Making as an opportunity for classroom assessment: Canadian maker educators' views on assessment" (International Journal of Child-Computer Interaction 39, (2024), 100631) https://doi.org/10.1016/j.ijcci.2023.100631 Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work in this paper.

DOIs of original article: https://doi.org/10.1016/j.ijcci.2023.100622, https://doi.org/10.1016/j.ijcci.2023.100621, https://doi.org/10.1016/j.ijcci.2023.100612, https://doi.org/10.1016/j.ijcci.2018.04.004, https://doi.org/10.1016/j.ijcci.2023.100574, https://doi.org/10.1016/j.ijcci.2018.04.003, https://doi.org/10.1016/j.ijcci.2024.100646, https://doi.org/10.1016/j.ijcci.2024.100634, https://doi.org/10.1016/j.ijcci.2023.100631, https://doi.org/10.1016/j.ijcci.2018.04.005, https://doi.org/10.1016/j.ijcci.2023.100632, https://doi.org/10.1016/j.ijcci.2023.100596, https://doi.org/10.1016/j.ijcci.2023.100594, https://doi.org/10.1016/j.ijcci.2023.100618, https://doi.org/10.1016/j.ijcci.2023.100583, https://doi.org/10.1016/j.ijcci.2023.100581, https://doi.org/10.1016/j.ijcci.2023.100578, https://doi.org/10.1016/j.ijcci.2024.100651, https://doi.org/10.1016/j.ijcci.2024.100635, https://doi.org/10.1016/j.ijcci.2023.100620, https://doi.org/10.1016/j.ijcci.2023.100601, https://doi.org/10.1016/j.ijcci.2023.100575, https://doi.org/10.1016/j.ijcci.2024.100650, https://doi.org/10.1016/j.ijcci.2023.100614, https://doi.org/10.1016/j.ijcci.2023.100591, https://doi.org/10.1016/j.ijcci.2023.100576, https://doi.org/10.1016/j.ijcci.2023.100566, https://doi.org/10.1016/j.ijcci.2023.100564, https://doi.org/10.1016/j.ijcci.2023.100602, https://doi.org/10.1016/j.ijcci.2022.100563.

https://doi.org/10.1016/j.ijcci.2024.100677

Available online 22 July 2024

""Hey, Alexa" "Hey, Siri", "OK Google" …." exploring teenagers' interaction with artificial intelligence (AI)-enabled voice assistants during the COVID-19 pandemic" (International Journal of Child-Computer Interaction 38, (2023), 100622) https://doi.org/10.1016/j.ijcci.2023.100622 Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work in this paper.

"Towards culturally sustaining design: Centering community's voices for learning through Participatory Design" (International Journal of Child-Computer Interaction 39, (2024), 100621) https://doi.org/10.1016/j.ijcci.2023.100621 Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work in this paper.

"Stimulating preschoolers' early literacy development using educational technology: A systematic literature review" (International Journal of Child-Computer Interaction 39, (2024), 100620) https://doi.org/10.1016/j.ijcci.2023.100620 Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work in this paper.

"AMBY: A development environment for youth to create conversational agents" (International Journal of Child-Computer Interaction 38, (2023), 100618) https://doi.org/10.1016/j.ijcci.2023.100618 Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work in this paper.

"Fast mapping in word-learning: A case study on the humanoid social robots' impacts on Children's performance" (International Journal of Child-Computer Interaction 38, (2023), 100614) https://doi.org/10.1016/j.ijcci.2023.100614 Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work in this paper.

"Shimpai Muyou: Confronting Child Suicides in Japan through a Serious Game & School-based Intervention" (International Journal of Child-Computer Interaction 38, (2023), 100612) https://doi.org/10.1016/j.ijcci.2023.100612 Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work in this paper.

"Making computing visible & tangible: A paper-based computing toolkit for codesigning inclusive computing education activities" (International Journal of Child-Computer Interaction 38, (2023), 100602) https://doi.org/10.1016/j.ijcci.2023.100602 Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work in this paper.

"ScratchJr design in practice: Low floor, high ceiling" (International Journal of Child-Computer Interaction 37, (2023), 100601 https://doi.org/10.1016/j.ijcci.2023.100601 Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work in this paper.

"A codesign study exploring needs, strategies, and opportunities for digital health platforms to address pandemic-related impacts on children and families" (International Journal of Child-Computer Interaction 37, (2023), 100596) https://doi.org/10.1016/j.ijcci.2023.100596 Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work in this paper.

"Child-home interaction: Design and usability evaluation of a game-based end-user development for children" (International Journal of Child-Computer Interaction 37, (2023), 100594) https://doi.org/10.1016/j.ijcci.2023.100594 Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work in this paper.

"From endpoints to trading zones: Multi-directional exchange for computational empowerment in computer science education" (International Journal of Child-Computer Interaction 37, (2023), 100591) https://doi.org/10.1016/j.ijcci.2023.100591 Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work in this paper.

"Making a Makerspace for children: A mixed-methods study in Chinese kindergartens" (International Journal of Child-Computer Interaction 36, (2023), 100583) https://doi.org/10.1016/j.ijcci.2023.100583 Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work in this paper.

"How Does Constructive Feedback in an Educational Game Sound to Children?" (International Journal of Child-Computer Interaction 36, (2023), 100581) https://doi.org/10.1016/j.ijcci.2023.100581 Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work in this paper.

"Assessment of learning in child–computer interaction research: A semi-systematic literature review" (International Journal of Child-Computer Interaction 36, (2023), 100578) https://doi.org/10.1016/j.ijcci.2023.100578 Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work in this paper.

"The situated power of computational empowerment" (International Journal of Child-Computer Interaction 36, (2023), 100576) https://doi.org/10.1016/j.ijcci.2023.100576 Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work in this paper.

"Students' achievement goals: Goal approximation, engagement, and emotions in co-design activities and product" (International Journal of Child-Computer Interaction 36, (2023), 100575) https://doi.org/10.1016/j.ijcci.2023.100575 Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work in this paper.

"Investigating primary school children's embodied expression of programming concepts" (International Journal of Child-Computer Interaction 36, (2023), 100574) https://doi.org/10.1016/j.ijcci.2023.100574 Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work in this paper.

"Systematic mapping of computational thinking in preschool children" (International Journal of Child-Computer Interaction 36, (2023), 100566) https://doi.org/10.1016/j.ijcci.2023.100566 Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work in this paper.

"Digital Media and Bildung—In memoriam of Heidi Schelhowe" (International Journal of Child-Computer Interaction 36, (2023), 100564) https://doi.org/10.1016/j.ijcci.2023.100564 Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work in this paper.

"Effects of a computational thinking intervention program on executive functions in children aged 10 to 11" (International Journal of Child-Computer Interaction 35, (2023), 100563) https://doi.org/10.1016/j.ijcci.2022.100563 Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work in this paper.