LUMAT

Special Issue: Technology in Mathematics Education

Vol 11 No 3
2023

Photo by: Stefan Krause / License: LAL

Editorial: Special Issue "Technology in Mathematics Education"

Mika Koskenoja and Topi Törmä

Department of Mathematics and Statistics, University of Helsinki, Finland

It is important to get more knowledge about the impact of the use of technology on the learning of mathematics. Recent research shows that the use of digital equipment and software is related to better mathematical skills, but excessive use of digital resources may disturb concentration on learning. This special issue presents five papers on different areas of technology in mathematics education.

Keywords: mathematics education, technology in education

ARTICLE DETAILS

LUMAT Special Issue
Vol 11 No 3 (2023)

Correspondence:
mika.koskenoja@helsinki.fi

https://doi.org/10.31129/LUMAT.11.3.2141

1 Remarks on the use of technology in mathematics education, particularly in Finland

Technological tools play an important role in teaching and studying mathematics. Calculators have aided straightforward computations for many decades, but nowadays curricula more and more demand the use of modern computer programs in school mathematics. However, it is not clear how and at which stage this technology should be implemented in mathematics education. Also, what is the impact of the use of technology on the learning of mathematics? At the latest, the covid-19 pandemic showed that technology and digital teaching materials are vital in education and must be developed and studied further. In this special issue we present several interesting technology related mathematics education research articles.

Finnish national core curricula for basic and for general upper secondary education (Finnish National Agency for Education, 2020b; 2020c) emphasize digitalization and the use of technology at schools. Moreover, Finnish national core curricula for early childhood education and care and for pre-primary education (Finnish National Agency for Education, 2020a; 2021) express that day care and pre-primary education should prepare children to the digitalized world. However, one notable and distracting observation in just released PISA 2022 results (OECD, 2023) is that 41% of Finnish 9th graders feel that the use of digital resources disturbs their concentration on learning mathematics. This is clearly more than the international average 31%. From the positive side, moderate and controlled use of digital equipment seems to be related

Published by the University of Helsinki, Finland / LUMA Centre Finland | CC BY 4.0

to better mathematical skills. Anyhow, due to PISA 2022 results (OECD, 2023) mathematical ability of 15-year-old school pupils is deteriorating in almost all over the world.

The Finnish matriculation examination is organized digitally and the last exam to become digital was the mathematics exam in spring 2019 (Finnish Matriculation Examination Board, 2023). Students in Finland use computers heavily in their studies during the whole three-year period they spend in upper secondary school. This is comprehensible since learning materials are mainly digital and students are required to be fluent computer software users when they complete their studies and take part in the matriculation examination. However, the change to digital school environment has occurred so fast that it has been somewhat uncontrolled. Therefore, new research is needed to indicate strengths and weaknesses of digitalized education.

Recently Mertala et al. (2022) made a descriptive and critical analysis of highly cited educational technology articles. They observed that in high impact factor journals by major publishers it is easier to publish positive than critical findings of the use of technology in education. This is a bit alarming since politicians and education authorities base their decisions on current research. During the editorial process of this special issue, we noticed that articles primarily including critical findings were not even submitted.

2 Papers included in the current issue

This special issue contains five papers on different areas of technology in mathematics education. They provide various aspects on how the use of technological tools influences teaching and learning of mathematics.

In the first paper, Gladys Sunzuma reviews integration of digital technologies into teaching and learning of geometry at the secondary school level. Due to the systematic literature review the most used technologies were augmented reality and dynamic geometry software GeoGebra and Cabri. Most of the reviewed articles focused on the effectiveness of technological tools for learning geometry, minority focused on implementation and development of technological tools. Quantitative, qualitative, and mixed methods research approaches were all used, but the majority of reviewed articles used qualitative methods mainly for empirical studies.

Nadine Yılmaz examines the technology-enhanced statistical problem-solving task design assignments prepared by pre-service teachers who will be mathematics teachers in the future. Participants of the case study prepared 28 task design assignments which were analyzed in terms of the characteristics of the data and the need of data, for example, in terms of learning goals, features of the data, being interesting/suitable for the level of the students and using statistical problem solving. Statistical problem-solving components (pose, collect, analyze, interpret) of the tasks were described with three levels, and majority of the tasks were found to be at mid-level which requires reading the data, reading between the data and reading beyond the data.

Anneli Dyrvold and Ida Bergvall study validity of computer-based assessment in mathematics. Empirical data was collected in an eye-tracking analysis of grade nine students who were divided into two groups that received different instructions about how to work with mathematics items with five types of functions. The results revealed that students are not equally equipped for a computer-based assessment that depends on preparatory instructions for the dynamic functions. They conclude that students need to be comfortable and very familiar with using dynamic and interactive functions to ensure validity of tests.

In a related article, Dyrvold and Bergvall study how students interact with digital teaching material including dynamic and interactive elements supplementing the static parts of the material. The data for this study was collected using an eye-tracking analysis of grade nine students from four different schools. The students worked on five mathematics items each of which was designed in five versions with increasing interactivity and dynamism. The results of the study showed that the students spend more time and attention on dynamic mathematical content than static content which can be useful to know when designing digital teaching material.

Raimundo Elicer, Andreas Tamborg, Kajsa Bråting and Cecilia Kilhamn compare the integration of programming and computational thinking into Danish and Swedish elementary mathematics teaching resources. They analyzed 165 tasks from teaching modules developed for a Danish pilot project integrating technology comprehension into school subjects, as well as 390 tasks from Swedish mathematics textbooks. They found that the two countries had taken quite different approaches to integrating computational thinking into school mathematics. The Danish tasks included a lot of data and statistics whereas the Swedish tasks emphasized patterns, sequences and following stepwise instructions.

3 Future research

There is a huge need for new research which is focused on the learning of different topics of mathematics when various technological tools are used. Using computer software like GeoGebra in geometry differs significantly from arithmetic and manipulation of algebraic formulas. Therefore, one uniform scheme for the use of technology on learning and teaching of all mathematics cannot be optimal. However, without targeted recommendations in curricula it is straightforward for teachers to let students to use computers for solving all kinds of mathematical problems, even though learning certain topics, and solving problems would benefit using paper and pen.

References

Finnish Matriculation Examination Board (2023). Matriculation Examination Board. https://www.ylioppilastutkinto.fi/en

Finnish National Agency for Education (2020a). *National core curriculum for pre-primary education 2014, second edition*. Finnish National Agency for Education.

Finnish National Agency for Education (2020b). *National core curriculum for basic education 2014, third edition*. Finnish National Agency for Education.

Finnish National Agency for Education (2020c). *National core curriculum for general upper secondary education 2019*. Finnish National Agency for Education.

Finnish National Agency for Education (2021). *National core curriculum for early childhood education and care 2018*, second edition. Finnish National Agency for Education.

Mertala, P., Moens, E., & Teräs, M. (2022). Highly cited educational technology journal articles: A descriptive and critical analysis. *Learning, Media and Technology*, 1–14. https://doi.org/10.1080/17439884.2022.2141253

OECD (2023). *PISA 2022 results (Volume I): The state of learning and equity in education, PISA*. OECD Publishing.

Technology integration in geometry teaching and learning: A systematic review (2010–2022)

Gladys Sunzuma

Bindura University of Science Education, Zimbabwe

Technology advancement provides an opportunity for helping both teachers and students to solve and improve mathematics teaching and learning performances. This systematic review aims to add to the discussion through a comprehensive overview of the integration of digital technologies into the teaching and learning of geometry at the secondary school level. A systematic literature review was conducted following the preferred reporting items for systematic reviews and meta-analyses (PRISMA) guidelines, with a focus on publication trends, types of technologies used, types of contributions, learning domains and research methods. Twenty-nine articles published between 2010 and 2022 were searched from the ERIC database. The findings showed that the majority of the articles were published in 2015 and the technologies that were used were GeoGebra, augmented reality, computer animation package, video-based cooperative, graphing calculator, micromedia flash, Powtoon animation, learning management system, interactive whiteboard, digital simulations-applets, iPads and tablet. Most of the reviewed articles focused on the effectiveness of the technologies in geometry teaching and learning. The findings indicated that the majority of the reviewed articles used quantitative research methods followed by qualitative methods studies. It is suggested that other studies be conducted with other databases and focus on challenges of integrating technology into the teaching and learning of geometry.

ARTICLE DETAILS

LUMAT Special Issue
Vol 11 No 3 (2023), 1–18

Received 15 February 2023
Accepted 5 July 2023
Published 22 August 2023

Pages: 18
References: 62

Correspondence:
gsunzuma@gmail.com

https://doi.org/10.31129/LUMAT.11.3.1938

Keywords: geometry, teaching, learning, technology, systematic review

1 Introduction

Mathematics consists of several components such as statistics, algebra and geometry among others. Geometry is a vital component in mathematics that includes the nature and relation between points, lines, shapes and space. Geometry is the mathematics knowledge that involves the nature of shape and space, measurement, magnitude as well as the relations of dots, lines, corners and surfaces (Abd Rahim et al., 2018). Geometry is an exceptionally rich area of knowledge, not merely for its great diversity and assortment, but in addition for its practical applications such as visual presentations, computer animation, virtual reality, and medicine (in the area of medical imaging, which led to substantial new results in fields such as geometric tomography), robotics, geometric modelling (including design, modification and the manufacture of cars and aeroplanes, in the construction of buildings, etc.) and

Published by the University of Helsinki, Finland / LUMA Centre Finland | CC BY 4.0

computer-aided design (CAD) (Viseu et al., 2022). According to Jones (2000), several modern developments in mathematics are largely geometric, for instance, geometric algebra (a representational and computational system for geometry that is entirely distinct from algebraic geometry), mathematical visualisation (the art of transforming the symbolic into geometry) and work on dynamical systems (a discipline closely intertwined with the main areas of mathematics). Due to its multiplicity of applications, there is a need for schools to promote geometry learning (Septia et al., 2018). Some of the reasons for including geometry in the mathematics curriculum and teaching in schools are helping students to think visually, helps in solving problems in other mathematics-related fields, helps students who experience abstraction problems and that the world is built by form and space (Petrus et al., 2017). In many countries globally, the objective of including geometry in the school curricula is to enable students to develop skills of problem-solving, visualisation, intuition, critical thinking, perspective, conjecturing, logical argumentation, deductive reasoning as well as the ability to produce proof (Jones & Tzekaki, 2016; Kuzniak, 2018; Horsman, 2019). In addition, the purpose of teaching geometry in schools is that students can use visualization; have spatial abilities as well as geometry modelling skills to solve problems (NCTM, 2000).

However, it was noted that the desired objectives associated with teaching geometry could not be accomplished and the conceptual understanding of geometry concepts could not be developed (Gülburnu, 2022). Regardless of the importance and popularity of geometry, researchers (Sutiarso et al., 2018; Nursyahidah, 2016) noted many difficulties associated with its teaching and learning and most students experience difficulties in learning geometry. Furthermore, research has established that geometry is one of the components of mathematics that is abstract and complex that both teachers and students find difficult to teach and learn (Gambari et al, 2014). Amongst the causes of the student's difficulties with geometry are misconceptions of geometry (Sutiarso et al., 2018) as well as the abstract and conventional approach of teaching that makes students learn by heart without understanding the concepts (Bergstrom & Zhang, 2016; Abdul Hanid et al., 2022). Such approaches have contributed to poor achievement in geometry.

Mathematics educators have been constantly searching for innovative approaches to teach mathematics for understanding including improving students' achievement and performance (Mensah & Nabie, 2021). An innovative approach to mathematics teaching that motivates learning and promotes higher achievement as well as improves the performance of students is the integration of digital technologies in the teaching and learning process (Mensah & Nabie, 2021; Tay & Mensah-Wonkyi, 2018). Two major purposes of using digital technologies in mathematics teaching are supporting the organisation of the teacher's work such as formative and summative assessment of students and producing learning materials as well as support for new approaches to doing and representing mathematics (Clark-Wilson et al., 2020). According to Ayan and Isiksal Bostan (2016), the integration of digital technologies in mathematics teaching activities enables students to be actively involved and be in a position to solve complex problems. According to Klančar et al. (2019) using digital technologies in the teaching and learning process enables the designing of rich learning environments through the use of varied digital materials and digital support tools such as simulations, animations and applets. Such technologies support different methods of teaching, for instance, experimentation, simulation, modelling, and research including solving routine mathematical problems and non-routine problems (Klančar et al., 2019). Digital technologies augment the learning of mathematics by facilitating practical, problem-solving and collaborative methods of teaching and learning (Žakelj & Klančar, 2022). Given such benefits of integrating digital technologies into the teaching and learning of mathematics, this study is a systematic review of integrating technology into the teaching and learning of geometry.

2 Technology in geometry teaching and learning

The teaching and learning of geometry require students to be able to imagine, construct and understand the construction of shapes to relate them with associated facts (Praveen & Kwan Eu, 2013). Hence, digital technologies will help students in imagining, and making observations and facts (Praveen & Kwan Eu, 2013). Numerous digital technologies are available for the teaching and learning of geometry, for instance, Geometers Sketchpad, calculators, interactive whiteboards, and GeoGebra (Praveen & Kwan Eu, 2013). GeoGebra is a dynamic geometric software that amalgamates statistics, calculus, algebra, geometry, arithmetic, and spreadsheet elements into a solitary easy-to-use package that enables the learning and teaching of mathematics at various stages (Abebayehu & Hsiu-Ling, 2021).

Research has revealed that geometry concepts taught using computer-based technology result in improved student achievement as compared to the conventional approaches that rely on the use of textbooks (Christou et al., 2006; Abdul Hanid, 2022). The integration of technology in geometrical learning is crucial as it enables students to understand the geometry concepts' problem solving process, for instance, the use of various problem-solving approaches including reducing misconceptions to understand geometry concepts (Hwang et al., 2009). A study by Gutiérrez (1996) revealed students' ability to solve geometry problems using software that helped them to manipulate 3D Geometry object essentially for the visualisation and mental image. Students will be attentive and actively involved in geometry concepts taught through the use of technology (Hollebrands & Okumuş, 2018). In addition, digital technologies provide students with an opportunity to use the varied technology resources for the geometry content and solve any problem (Lee & Hollebrands, 2006).

Regarding the current systematic literature review about integrating digital technologies into the teaching and learning of mathematics, studies have been conducted. Mohamed et al. (2022), for instance, provides a systematic review of artificial intelligence in mathematics education. Zhong and Xia (2020) provide a stimulating learning experience with robotics in the learning of mathematics. Ahmad and Junaini (2020) focused on augmented reality in the teaching and learning of mathematics. A systematic review was also conducted by Abebayehu and Hsiu-Ling (2021) on the use of GeoGebra in the teaching and learning of mathematics. Even though these studies focused on digital technology integration into the teaching and learning of mathematics, they focused on mathematics in general. The current study focuses on the integration of digital technologies into particular mathematics topics geometry at secondary school level. To direct this systematic literature review, the current study addressed the following research questions:

1. What is the trend of articles on the integration of technologies into geometry teaching and learning from 2010 to 2022?
2. What are the leading technologies that have been integrated into the teaching and learning of geometry?
3. What are the types of contributions made by the articles in terms of implementation, development and effectiveness in geometry teaching and learning?
4. What are the learning domains in the teaching and learning of geometry?
5. What are the research methods used to study technology integration in geometry teaching and learning?

3 Methodology

A systematic search was conducted using the PRISMA specification that enables transparent and comprehensive reporting of systematic reviews (Page et al., 2021). Articles published in indexed journals are generally more systematically scrutinized such that they have a greater impact on the area of study (Duman et al., 2015). In this study, articles were searched from Education Resources Information Center (ERIC) because it is a chief source of high quality indexed academic journals (Akçayır & Akçayır, 2018). The search function was used and input the keywords "Geometry" or "Secondary level" or "Teaching and Learning" or "Technology" and "Information Communication Technology (ICT)."

The screening criteria excluded book chapters, books, conference proceedings, systematic review articles, or books. The study focused on English-language journal articles to avoid complex or uncertain translations. Journal articles published between the years 2010 and 2022 focusing on secondary school level geometry teaching and learning and technology integration were included. From the screening process, 873 articles were identified. To guarantee that all 873 articles fit the study's selection criteria and objectives, each article's title, abstract, methodology, results, and discussion were scrutinized. Ten articles were removed as they were duplicates. 812 articles were rejected because of the following reasons; they did not explain how technologies were integrated into the teaching and learning of geometry at the secondary school level, not written in English and were books and conferences. Another 32 articles were rejected because they focused on the teachers' use of technology only without the teaching and learning component. Finally, 29 articles were included in the final stage of the review process as shown in Figure 1. Thematic analysis was carried out to classify the themes related to the research trends and patterns in the study. Useful data was extracted from the 29 articles that were used to answer the research questions. The 29 articles used in this study were marked with an asterisk in the list of references.

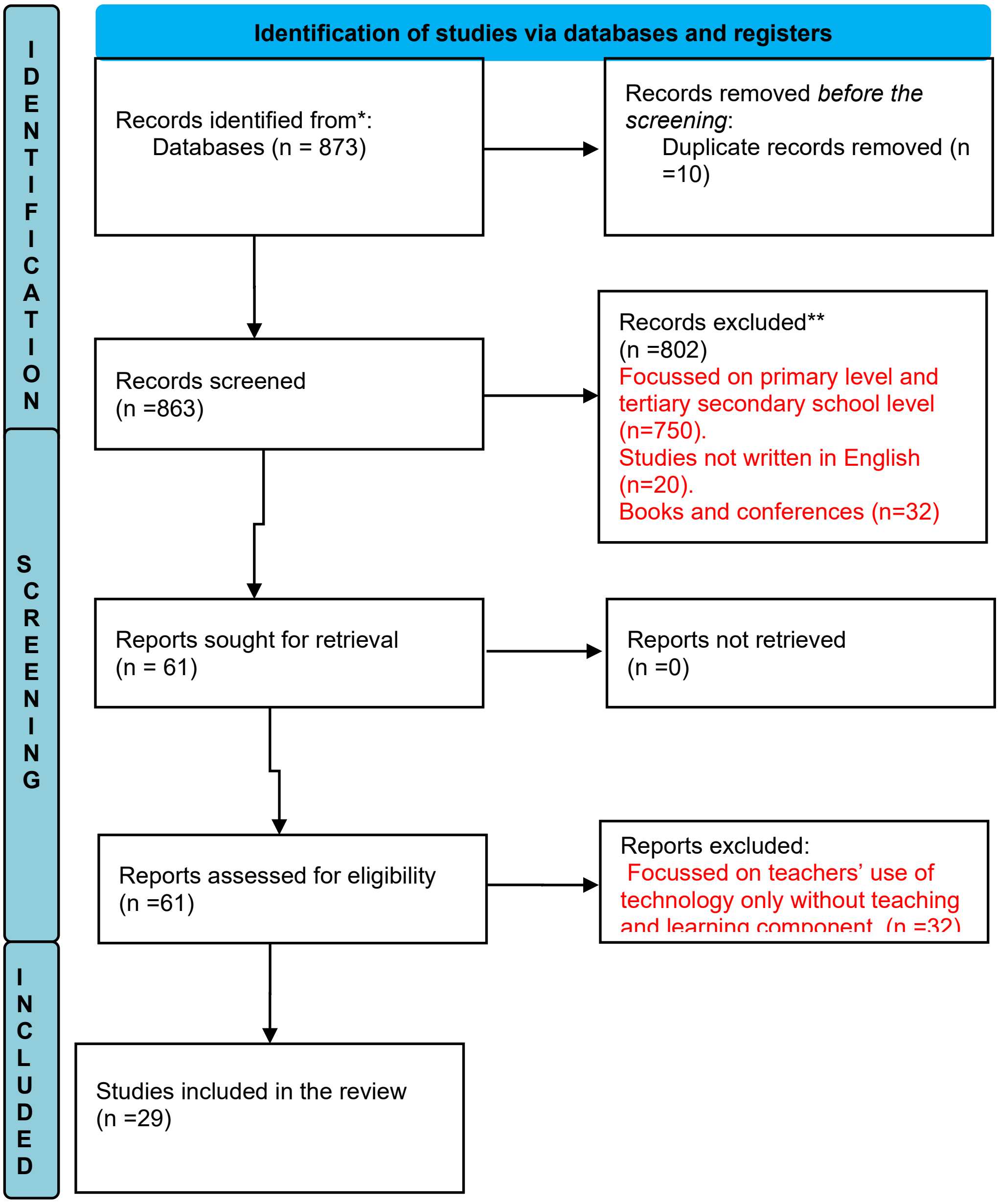


Figure 1. Systematic literature review procedure used in the study following concepts in Page, et al. (2021).

4 Results and Discussion

A search performed on the ERIC database resulted in 873 journal articles. Only 29 articles met the inclusion criteria. The findings of the systematic review are presented under the following themes: trends of article publication; type of technologies; type of contribution (development, implementation, and effectiveness), learning domain and research approaches used to carry out the studies.

4.1 Trends in article publication

The trend of the 29 published articles integrating different technologies into the teaching and learning of geometry from 2010 to 2022 is shown in Figure 2. To comprehend the development of the research, the 29 articles were classified based on the year of publication. There were no articles published in 2012 and 2017 (Figure 3). There is a gradual increase in articles from 2013 to 2015 with the highest number of seven articles recorded in 2015. The number of published articles remained relatively consistent from 2021 to 2022. The general trend shows that some researchers focus on the integration of technology into the teaching and learning of geometry. Such findings from the systematic review show that even though the articles were very few in terms of number some progress has been made in the integration of technology into the teaching and learning of geometry. Even though the trend is moving upwards and downwards there is evidence that the issue of technology integration in geometry teaching and learning is progressively becoming an area of focus that is getting numerous researchers' attention.

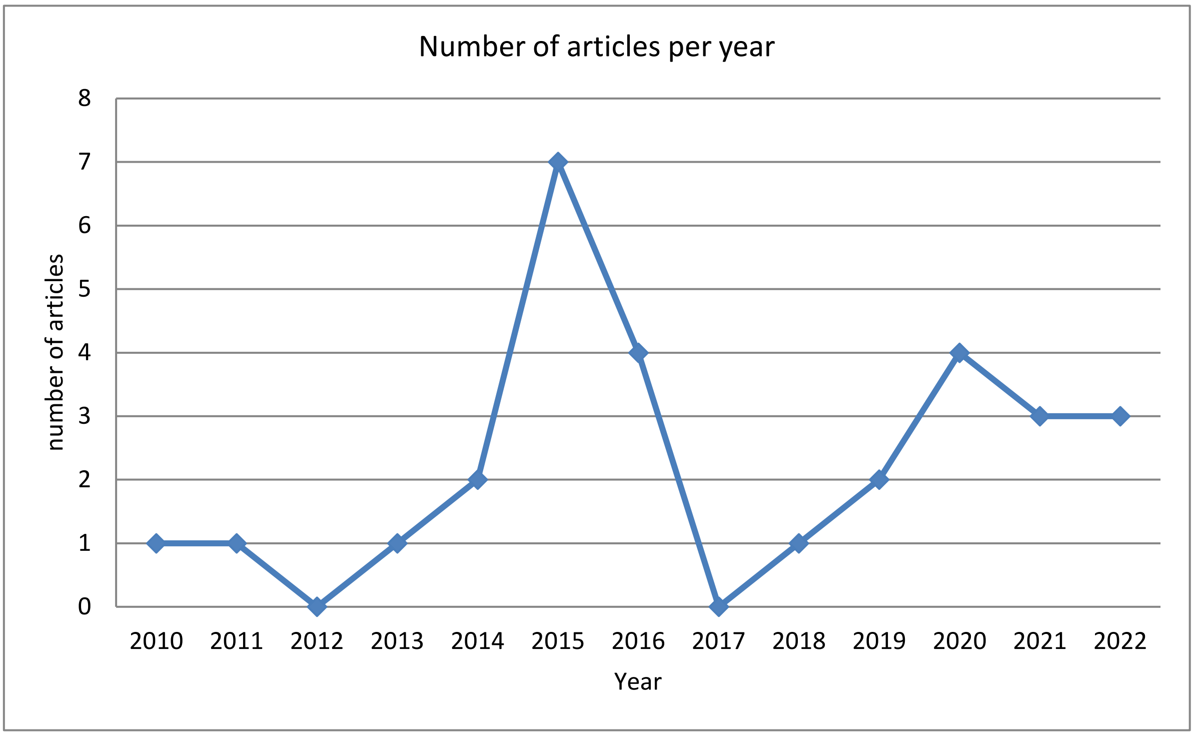


Figure 2. Number of articles per year.

4.2 Types of technologies used

Different types of technologies were used in the teaching and learning of geometry as shown in Figure 3. The majority of the studies used dynamic geometry software with GeoGebra having a total of 10 articles, Cabri had two articles, multiuser dynamic geometry, dynamic geometry general, and dynamic geometry of sketchpad each having one article. The findings of the current review are in line with Abebayehu and Hsiu-Ling (2021) who found out that GeoGebra is widely used in the teaching and learning of geometry. Geometry is one of the most frequent mathematics topics that integrate GeoGebra because of its potential to visualize abstract and difficult concepts through many representations. Representations help students to understand and make associations between geometry concepts. Visualization is not merely pertinent for illustrative purposes but is as well acknowledged as an essential component of problem solving, reasoning and even proofs (Abebayehu & Hsiu-Ling, 2021).

Augmented Reality had the second largest number of four articles. Although augmented reality has the second largest number of articles, a study by Ahmad and Junaini (2020) showed its wide use in the teaching and learning of geometry. Augmented reality in geometry teaching and learning provides students with an interactive learning environment, increased understanding and retention as well as enhanced visualization (Ahmad & Junaini, 2020). Computer animation package, video-based cooperative, graphing calculator, micromedia flash, Powtoon animation, learning management system, interactive whiteboard, digital simulations-applets, iPads and tablet device each had only one article. The limited use of such digital technologies could be due to the lack of adequate and ample training for teachers as observed by Dockendorff and Solar (2018) who reported that a lot of teachers are inadequately prepared to incorporate digital technology into the mathematics curriculum.

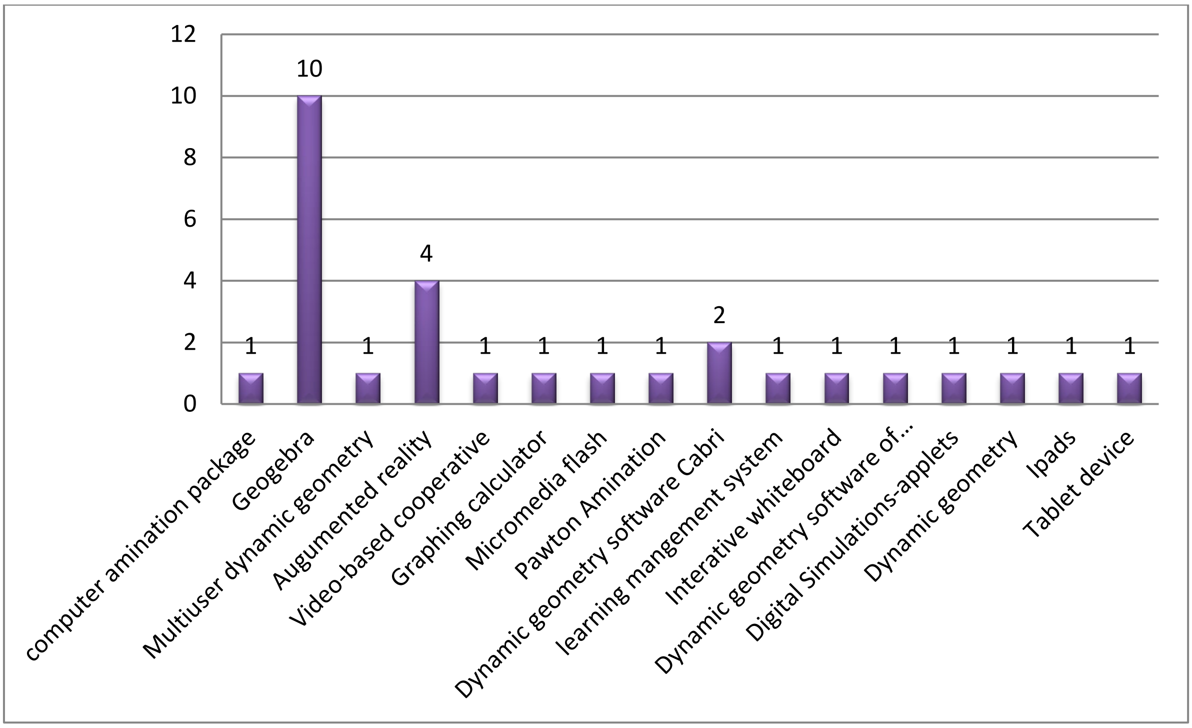


Figure 3. Types of technologies used.

4.3 Types of contributions (Implementation, development, and effectiveness)

Figure 4 shows the distribution of the types of contributions made by the articles to the teaching and learning of geometry in terms of implementation, development and effectiveness of technological tools for geometry learning. Twenty-one articles (73%) focused on the effectiveness of technological tools in the teaching and learning of geometry (Gambari et al. 2014; Doğan & İçel, 2011; Ibili et al. 2020; Praveen & Kwan Eu 2013; Gambari et al. 2016; Diaz-Nunja et al. 2018; Kandemir & Demirbag, 2019; et al. 2019; Yani & Rosma, 2020; Mailizar., & Johar, 2021; Brito et al. 2021; Shaame et al. 2020; Perry & Steck, 2015; Gómez-Chacón et al. 2016; Lin et al. 2014; Samur Turk & Akyüz, 2016; Abdul Hanid et al. 2022; Viseu et al. 2022; Lin et al. 2015; Duroisin et al. 2015). The greatest number of articles focused on the effectiveness of the technological tools as shown in Figure 4. Five articles (17%) focused on the implementation of technological tools in the teaching and learning of geometry (Ng & Sinclair, 2015; Lin et al. 2015; Prasad, 2016; Komatsu & Jones, 2020; Gülburnu, 2022). Only three articles (10%) focused on development. Akmalia et al. (2021) developed powtoon animation. Sherman and Cayton (2015) developed a framework for teaching geometry using technology. Baccaglini-Frank and Mariotti (2010) developed a dragging model for generating conjectures in dynamic geometry. The finding of this review disagrees with earlier findings by Ahmad and Junaini (2020) where the findings showed that the major contribution of the articles reviewed was the development of apps. In view of the fact that most of the articles in the current review focused on effectiveness, additional research on technology integration into the teaching and learning of geometry other than effectiveness should be conducted even more.

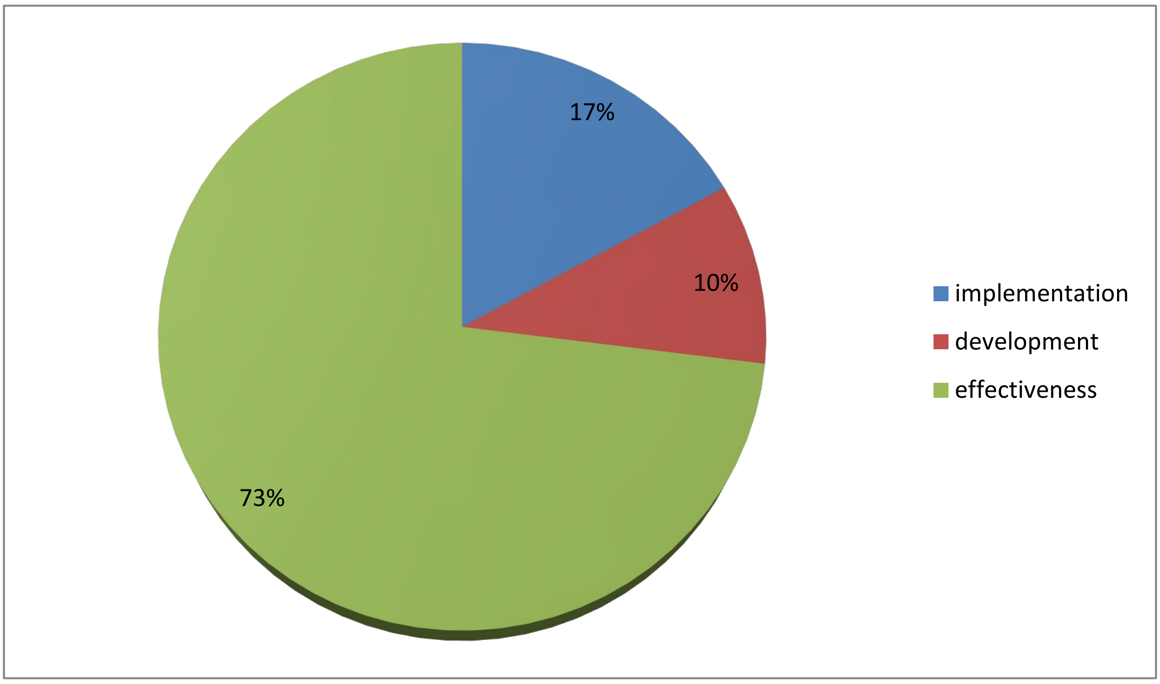


Figure 4. Types of contributions.

4.4 Learning domains

Technology integration into the teaching and learning of geometry had been classified based on Bloom's Revised Taxonomy of learning domains which includes the cognitive domain, the affective domain, and the psychomotor domain (Krathwohl, 2002). The cognitive learning domain which includes the component of obtaining knowledge from learning as well as the development of intellectual capabilities through low to higher-order learning such as problem solving and learning performance had ten articles (Gambari et al. 2014; Doğan & İçel, 2011; Lin et al. 2015; Praveen & Kwan Eu, 2013; Gambari et al. 2016; Kandemir & Demirbag, 2019; Adelabu et al. 2019; Kaushal Kumar & Chun-Yen, 2015; Shaame et al. 2020; Samur Turk & Akyüz, 2016). The articles revealed positive effects on student learning and achievement. They found an increased achievement in geometry after using various technologies. The study by Gómez-Chacón et al. (2016) shows that the use of dynamic geometry affords students greater intellectual independence in geometrical work, whilst Ibili et al. (2020) found out found that geometry teaching supported by Augmented Reality increased the students' 3D thinking skills. A study by Fukawa-Connelly and Silverman (2015) focused on the development of mathematical argumentation in an unmoderated, asynchronous multi-user dynamic geometry environment. The study showed that the students made progressively more in-depth and mathematical descriptions of the data, developed more conceptual warrants, as well as progressively behaved as if giving reasons was normative in the discussion.

The affective learning domain involves students' feelings about learning, for example, motivation and learning perceptions. Seven articles had issues to do with the affective domain. The integration of technology into the teaching and learning of geometry improves students' motivation as mentioned in one study. The study by Doğan and İçel (2011) has shown that the use of GeoGebra improves students' motivation with a positive impact. Five studies (Samur Turk & Akyüz, 2016; Gómez-Chacón et al. 2016; Gülburnu, 2022; Duroisin et al. 2015; Lin et al. 2015) found the benefits of various technologies in developing students' positive attitudes towards geometry learning. The study by Perry and Steck (2015) assessed the effect of integrating iPads in geometry teaching on student engagement, self-efficacy, and meta-cognitive self-regulation. The finding showed that the students who used the iPad experienced higher levels of off-task behaviours and similar levels of self-efficacy and meta-cognitive self-regulation as compared to the group that did not use the iPad.

The psychomotor learning domain involves the manipulation or motor skill area of learning such as spatial skills. The integration of technologies into the teaching and learning of geometry enhances the spatial ability and visualization skills. The study by Yani and Rosma (2020) showed an improvement in students' spatial ability and visualization skills after the use of the macromedia flash. Abdul Hanid et al. (2022) stated that Augmented Reality enhances students' visualization skills. Meanwhile, technology has made geometry learning more interactive. Three studies explained this benefit. Gómez-Chacón et al. (2016) stated that dynamic geometry software affords interaction with the context that impacts learning opportunities in geometric proofs; whilst Gülburnu (2022) was of the idea that Cabri 3D encourages interaction through facilitating drawings and measurements. Duroisin et al. (2015) stated that the use of the interactive whiteboard encourages interactions between the students and has a positive effect on the efficiency of the learning sequence itself.

A study by Gülburnu (2022) showed that Cabri 3D encourages the association of geometric knowledge about solids volume measurement with daily life by contributing to conceptual and permanent learning. GeoGebra offers students an opportunity to experiment and explore that result in improved results (Viseu et al. 2022).

4.5 Research Approaches used in the studies

Different research methods were employed in the 29 articles. The research findings show that only three research approaches which are quantitative, qualitative and mixed methods were used in the 29 articles as shown in Figure 5. The analysis revealed that the majority of the studies reviewed (62%, n=18), (Gambari et al. 2014; Doğan & İçel, 2011; Ibili et al. 2020; Praveen & Kwan Eu 2013; Gambari et al. 2016; Diaz-Nunja et al. 2018; Kandemir & Demirbag, 2019; Adelabu et al. 2019; Yani & Rosma, 2020; Mailizar., & Johar, 2021; Akmalia et al. 2021; Brito et al. 2021; Shaame et al. 2020; Perry & Steck, 2015; Gómez-Chacón et al. 2016; Lin et al. 2014; Samur Turk & Akyüz, 2016; Abdul Hanid et al. 2022) used quantitative research methods. The quantitative methods were mainly empirical studies. According to Yang et al. (2019), an empirical study is carried out to examine the cause-and-effect relationship between independent and dependent variables under conditions of apt control hence it is regarded as the most scientific method among all the experimental research. The quantitative approach was mainly chosen as it put more emphasis on the objective measurement and analysis of numerical or statistical, data collected through tests, surveys and questionnaires.

Seven studies (24%) of the reviewed studies used qualitative research methods (Ng & Sinclair, 2015; Fukawa-Connelly & Silverman, 2015; Gülburnu, 2022; Sherman & Cayton, 2015; Prasad, 2016; Baccaglini-Fran &, Mariotti, 2010; Komatsu & Jones, 2020). Qualitative research methods involve collecting and analyzing non-numerical data with the purpose of a better understanding of concepts, views, or experiences. Case studies are employed to examine a phenomenon in-depth as well as to understand particular situations and provide an in-depth analysis (Olsson, 2018). For example, Gülburnu (2022) employed a case study that enabled the researcher to investigate students' views on geometry teaching through the use of the three-dimensional dynamic geometry software Cabri 3D.

Although mixed methods incorporate the benefits of both quantitative and qualitative methods, the current systematic review showed that only four studies (14%) used the mixed method research (Viseu et al. 2022; Kandemir & Demirbag, 2019; Duroisin et al. 2015; Lin et al. 2015). Mixed methods are of use in understanding inconsistencies between quantitative and qualitative findings and they enhance the problem by comparing the findings. For example, in a study by Viseu et al. (2022), the quantitative method focused fundamentally on the characteristics of the student's answers with regard to their level of correctness, whilst the qualitative method focused as well on the students' answers but the intention to analyse the reasons beyond such answers. Mixed method research fosters intellectual interaction and flexibility since researchers would expand the distribution of data on technology integration into the teaching and learning of geometry.

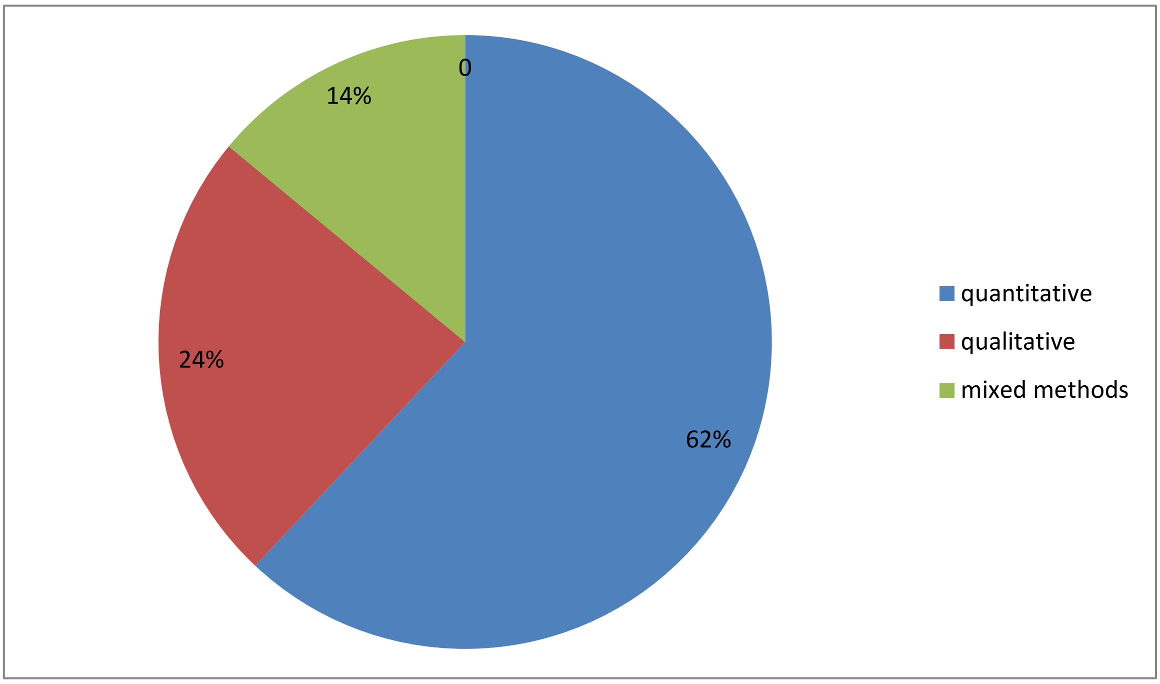


Figure 5. Research approaches used.

5 Conclusion and suggestions

The teaching and learning of geometry are challenging because most of the concepts are abstract. One effective method of improving geometry learning is through technology integration into the teaching and learning process. Students can investigate, solve, and explain geometrical concepts in different forms in a technology-rich environment. The current review assists to understand a systematic and comprehensive examination over the last twelve years of research in technology integration into the teaching and learning of geometry at the secondary school level. In addition, an updated analysis was provided that reveals learning and technological requirements for further studies to be conducted in the future. Restricted by the scope of the current review findings were classified into themes such as publication trends, types of technologies used, types of contributions, learning domains and research methods. Technology integration into the teaching and learning of geometry helps to ease the process of learning.

The systematic review revealed that GeoGebra is widely used in the teaching and learning of geometry followed by augmented reality. This review also found that the most observed contribution of the articles is the effectiveness of technology integration into the teaching and learning process. It is important to be acquainted with the extent of the effectiveness of technology integration into the teaching and learning of geometry to enable its wide application in the future if results in positive effectiveness. In learning domains, the cognitive domain focused on students' learning achievement and 3D thinking skills. The affective domain was based on assessing student engagement, self-efficacy, meta-cognitive, motivation as well as perceptions. The psychomotor domain focused on students' visualization skills, spatial ability and interactive learning in geometry. The review reveals a clear outline of the often-used research methods employed in the articles incorporated in this study. For example, the quantitative research method was the most commonly used approach in articles on technology integration in the teaching and learning of geometry, whilst mixed methods had the least number of articles.

6 Limitation and Implications for Future Studies

The search terms used in the methodology are a limitation of this study. Only the articles published in ERIC database were included in this study. Therefore, further studies could be conducted using other databases. In addition, there is a need to carry out more research on the negative side effects of using technology in the teaching and learning of geometry.

References

Abebayehu Y., & Hsiu-Ling, C. (2021). GeoGebra in mathematics education: a systematic review of journal articles published from 2010 to 2020. *Interactive Learning Environments*, DOI: 10.1080/10494820.2021.2016861

Ahmad, N. I. N., & Junaini, S. N. (2020). Augmented Reality for Learning Mathematics: A Systematic Literature Review. *International Journal of Emergency Technology learning, 15*(16), 106–122

Abd Rahim, F., Ujang, N., & Said, M. T. (2018). Geometri dan peranannya dalam reka bentuk ban dar Islamik. *Malaysian Journal of Society and Space, 14*(2), 82–96. https://doi.org/10.17576/geo-2018--1402-07

*Abdul Hanid, M., Mohamad Said, M., & Yahaya, N. (2022). Effects of augmented reality application integration with computational thinking in geometry topics. *Education Information Technolo*gy 27, 9485–9521. https://doi.org/10.1007/s10639-022-10994-w

*Adelabu, F. M, Makgato, M., and Ramaligela, M.S., (2019). The Importance of Dynamic Geometry Computer Software on Learners' Performance in Geometry. *The Electronic Journal of e-Learning*, *17*(1), 52–63

Ayan, R., & Isiksal Bostan, M. (2016). Middle School Students? Reasoning in Nonlinear Proportional Problems in Geometry. *International Journal of Science and Mathematics Education, 16,* 1–16. https://doi.org/10.1007/s10763-016-9777-z

Akçayır, G., & Akçayır, M. (2018). The flipped classroom: A review of its advantages and challenges. *Computers & Education,* 126, 334–345. https://doi.org/10.1016/j.compedu.2018.07.021

*Akmalia, R., Fajriana; R., Nufus, H., & Wulandari, W. (2021). Development of powtoon animation learning media in improving understanding of mathematical concept. *Malikussaleh Journal of Mathematics Learning*, *4*(2), 105–116.

*Baccaglini-Frank, A., & Mariotti, M.A. (2010). Generating Conjectures in Dynamic Geometry: The Maintaining Dragging Model. *International Journal Computer Mathematics Learning 15*, 225–253. https://doi.org/10.1007/s10758-010-9169-3

Bergstrom, C., & Zhang, D. (2016). Geometry interventions for K-12 students with and without disabilities: A research synthesis. *International Journal of Educational Research, 80*, 134–154. https://doi.org/10.1016/j.ijer.2016.04.004

*Brito, L. P., Almeida, L. S. & Osório, A. J. (2021). Seeing in believing: impact of digital simulation pedagogical use in spatial geometry classes. *International Journal of Technology in Teaching and Learning, 17*(2), 109–123

Christou, C., Pittalis, M., Mousoulides, N., & Jones, K. (2006). Developing the 3DMath dynamic geometry software: theoretical perspectives on design. *International Journal, 13*, 168–174. http://eprints.soton.ac.uk/42114/

Clark-Wilson, A., Robutti, O., & Thomas, M. (2020). Teaching with digital technology. *ZDM Mathematics Education, 52*(7), 1223–1242. https://doi.org/10.1007/s11858-020-01196-0

*Diaz-Nunja, L., Rodríguez-Sosa, J., & Lingán, S.K. (2018). Teaching of Geometry with the GeoGebra Software in High School Students of an Educational Institution in Lima, *Propósitos y Representaciones 6*(2), 217–251. Doi: http://dx.doi.org/10.20511/pyr2018.v6n2.251

Dockendorff, M., & Solar, H. (2018). ICT integration in mathematics initial teacher training and its impact on visualization: The case of GeoGebra. *International Journal of Mathematical Education in Science and Technology, 49*(1), 66–84. https://doi.org/10.1080/0020739x.2017.1341060

*Doğan, M. İçel, R. (2011). The role of dynamic geometry software in the process of learning: GeoGebra example about triangles *International Journal of Human Sciences*, *8*(1). Available: http://www.InsanBilimleri.com/En

*Duroisin, N., Temperman, G., & De Lièvre, B (2015). Restrict or Share the Use of the Interactive Whiteboard? The Consequences on the Perception, the Learning Processes and the Performance of Students within a Learning Sequence on Dynamic Geometry. *The Turkish Online Journal of Educational Technology, 14*(2), 144–154

Duman, G., Orhon, G., & Gedik, N. (2015). Research trends in mobile assisted language learning from 2000 to 2012. *ReCALL, 27*(2), 197–216. https://doi.org/10.1017/S0958344014000287

*Fukawa-Connelly, T., & Silverman, J. (2015). The development of mathematical argumentation in an unmoderated, asynchronous multi-User dynamic geometry environment. *Contemporary Issues in Technology and Teacher Education, 15*(4), 445–488.

*Gambari, A.I., Falode, C.O., & Adegbenro, D. (2014). Effectiveness of Computer Animation and Geometrical Instructional Model on Mathematics Achievement and Retention among Junior Secondary School Students. *European Journal of Science and Mathematics Education, 2*, 127–146.

*Gambari, A.I., Shittu, A. T., Daramola, F. O., & James, M. (2016). Effects Of Video-Based Cooperative, Competitive And Individualized Instructional Strategies On The Performance Of Senior Secondary Schools Students In Geometry. *Malaysian Online Journal of Educational Sciences, 4*(4), 31–47

*Gómez-Chacón, I.M., Romero Albaladejo, I.M. & del Mar García López, M. (2016). Zig-zagging in geometrical reasoning in technological collaborative environments: a Mathematical Working Space-framed study concerning cognition and affect. *ZDM Mathematics Education* 48, 909–924 https://doi.org/10.1007/s11858-016-0755-2

*Gülburnu, M. (2022). Secondary School Students' Views on Geometry Teaching via Three-Dimensional Dynamic Geometry Software Cabri 3D: Solid Volume Measurement. *International Journal of Curriculum and Instruction*, *14*(1), 1088–1105

Gutiérrez, A. (1996). Visualization in 3-dimensionalgeometry: In search of a framework. In L. Puig, and A. Gutiérrez (Eds.), *Proceedings of the 20th PME Conference* (pp. 56-79).

Hollebrands, K.F., & Okumuş, S. (2018). Secondary mathematics teachers' instrumental integration in technology-rich geometry classrooms. *The Journal of Mathematical Behavior, 49*, 82–94.

Horsman, R., (2019). International perspectives on the teaching and learning of geometry in secondary schools. *Research in Mathematics Education, 21*(1), 96–100. https://doi.org/10.1080/14794802.2018.1531055

Hwang, W.-Y., Su, J.-H., Huang, Y.-M., & Dong, J.-J. (2009). A Study of Multi-Representation of Geometry Problem solving with virtual manipulatives and whiteboard system. *Educational Technology & Society, 12* (3), 229–247.

*İbili, E., Çat, M., Resnyansky, D., Şahin, S., & Billinghurst, M. (2020). An assessment of geometry teaching supported with augmented reality teaching materials to enhance students' 3D geometry thinking skills, *International Journal of Mathematical Education in Science and Technology*, *51*(2), 224–246

Jones, K. (2000), Critical Issues in the Design of the Geometry Curriculum. In Bill Barton (Ed). *Readings in mathematics education* (pp. 75-90). Auckland, New Zealand.

Jones, K. & Tzekaki, M. (2016). Research on the teaching and learning of geometry. In A Gutiérrez, GC Leder & P Boero (eds). The second handbook of research on the psychology of mathematics education: The journey continues. Rotterdam, The Netherlands

*Kandemir, M.A., & Demirbag Keskin, P. (2019). Effect of graphing calculator program supported problem solving instruction on mathematical achievement and attitude. *International Journal of Research in Education and Science, 5*(1), 203–223.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory Into Practice, 41*(4), 212–218. https://doi.org/10.1207/s15430421tip4104_2

Kuzniak, A. (2018). Thinking about the teaching of geometry through the lens of the theory of geometric working spaces. In P Herbst, U Cheah, P Richard & K Jones (eds). International perspectives on the teaching and learning of 12 Smith, Julie, Gierdien geometry in secondary schools. Cham, Switzerland: Springer. https://doi.org/10.1007/978- 3-319-77476-3_2

*Kaushal, K., & Chun-Yen, C. (2015). Incorporating GeoGebra into Geometry Learning - A Lesson from India. *EURASIA Journal of Mathematics, Science & Technology Education*, *11*(1), 77–86

Klančar, A., Cotič, M., & Žakelj, A. (2019). Učenje in poučevanje geometrije z uporabo informacijsko-komunikacijske tehnologije v osnovni šoli [Learning and teaching geometry using ICT in elementary school]. Založba Univerze na Primorskem. https://doi.org/10.26493/978-961-7055-63-4

*Komatsu, K., & Jones, K. (2020). Interplay between Paper-and-Pencil Activity and Dynamic-Geometry-Environment Use during Generalisation and Proving. *Digital Experiences in Mathematics Education, 2020*(6), 123–143. https://doi.org/10.1007/s40751-020-00067-3

Lee, H. S., & Hollebrands, K. F. (2006). Students' use of technological features while solving a mathematics problem. *Journal of Mathematical Behavior, 25*(3), 252–266.

*Lin, H, K., Chen, M., & Chang, K. (2015). Assessing the effectiveness of learning solid geometry by using an augmented reality assisted learning system. *Interactive Learning Environments, 23*(6), 799–810.

*Lin, J.R. H., & Lin, S.S.J. (2014). Cognitive load for configuration comprehension in computer-supported geometry problem solving: An eye movement perspective. *International Journal of Science and Mathematics Education 2014*(12), 605–627 https://doi.org/10.1007/s10763-013-9479-8

*Mailizar., & Johar, R. (2021). Examining Students' Intention to Use Augmented Reality in a Project-Based Geometry Learning Environment. *International Journal of Instruction, 14*(2), 773–790.

Mohamed, M. Z., Hidayat, R., Suhaizi, N. N., Sabri, N. M., Mahmud, M. K. H., & Baharuddin, S. N. (2022). Artificial intelligence in mathematics education: A systematic literature review. *International Electronic Journal of Mathematics Education, 17*(3),

Mensah, J. Y., & Nabie, M. J. (2021). The effect of PowerPoint instruction on high school students' achievement and motivation to learn geometry. *International Journal of Technology in Education (IJTE), 4*(3), 331–350. https://doi.org/10.46328/ijte.55

NCTM. (2000). Principles and Standards for School Mathematics. Reston: National Council of Teachers of Mathematics.

*Ng, O., & Sinclair, N. (2015). Area Without Numbers": Using Touchscreen Dynamic Geometry to Reason About Shape. *Canadian Journal of Science, Mathematics and technology Education. 15*(1), 84–101

Nursyahidah, F., Saputro, B, A., & Prayito, M. (2016). Kemampuan Penalaran Matematis Siswa SMP dalam Belajar Garis dan Sudut dengan GeoGebra. *Suska Journal of Mathematics Education, 2*(1), 13–19

Olsson, J. (2018). The contribution of reasoning to the utilization of feedback from software when solving mathematical problems. *International Journal of Science and Mathematics Education, 16*(4), 715–735. https://doi.org/10.1007/s10763-016-9795-x

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., Stewart A.L., Thomas, J., Tricco, A.C., Welch. V. A., Whiting, P., & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ (Clinical research ed.), 372*, n71. https://doi.org/10.1136/bmj.n71

*Perry, D. R., & Steck, A. K. (2015). Increasing Student Engagement, Self-Efficacy, and Meta-Cognitive Self-Regulation in the High School Geometry Classroom: Do iPads Help? *Computers in the Schools, 32*(2), 122–143, DOI: 10.1080/07380569.2015.1036650

Petrus, Z., Karmila., & Riady, A. (2017). Deskripsi Kemampuan Geometri Siswa SMP Berdasarkan Teori Van Hiele. *Pedagogy, 2*(1), 145–160

*Praveen, S., & Kwan Eu, L. (2013). Effectiveness of Using GeoGebra on Students' Understanding in Learning Circles**.** *Malaysian Online Journal of Educational Technology*, *1*(4) 1–11

*Prasad, P. V. (2016). Leveraging Interactive Geometry Software to Prompt Discussion. *Mathematics Teaching in the Middle School*, *22*(4), 226–233. https://doi.org/10.5951/mathteacmiddscho.22.4.0226

*Samur Turk, H., & Akyüz, D. (2016). The Effects of Using Dynamic Geometry on Eighth Grade Students' Achievement and Attitude towards Triangles. *International Journal for Technology in Mathematics Education*, 95–102. https://hdl.handle.net/11511/35759

Septia, T., Prahmana, R. C. I., Pebrianto, & Wahyu, R. (2018). Improving Students Spatial Reasoning with Course LAB. *Journal on Mathematics Education, 9*(2), 327–336

*Shaame, A.A., Anatory, J., Osaki, K.M., & Mrutu, S.I. (2020). Exploring a Learning Management System as a Way to Improve Students' Understanding of Geometry in Secondary Schools. *Africa Education Review, 17*, 17–40.

*Sherman, M. F., & Cayton, C. (2015). Using appropriate tools strategically for instruction. *Mathematics Teacher, 109*(4), 306–310.

Sutiarso, S., Coesamin, M., & Nurhanurawati. (2018). The effect of various media scaffolding on increasing understanding of geometry concepts in elementary school students. *Journal on Mathematics Education, 9*(1), 95–102.

Tay, M. K, & Mensah-Wonkyi, T. (2018). Effect of using GeoGebra on senior high school students' performance in circle theorems. *African Journal of Educational Studies in Mathematics and Sciences, 14*, 1–17

*Viseu, F., Rocha, H., & Monteiro, J. M. (2022). Rethinking Digital Technology versus Paper and Pencil in 3D Geometry. *Journal of Learning for Development*, *9*(2), 267–278

*Yani, M., & Rosma, F. (2020). Improving Students' Spatial Ability by Using Macromedia Flash on Geometry Materials. *Malikussaleh Journal of Mathematics Learning, 3*(1), 18–22

Yang, Q.F., Lin, C. J., & Hwang, G.J. (2019). Research focuses and findings of flipping mathematics classes: A review of journal publications based on the technology-enhanced learning model. *Interactive Learning Environments, 29*(4), 1–34. https://doi.org/10.1080/10494820.2019.1602842

Žakelj, A., & Klančar, A. (2022). The role of visual representations in geometry learning. *European Journal of Educational Research, 11*(3), 1393–1411. https://doi.org/10.12973/eu-jer.11.3.1393

Zhong, B., & Xia, L. (2020). A systematic review on exploring the potential of educational robotics in mathematics education. *International Journal of Science and Mathematics Education, 18*(1), 79–101. https://doi.org/10.1007/

Examination of technology-enhanced statistical problem-solving tasks designed by pre-service teachers

Nadide Yılmaz

Karamanoglu Mehmetbey University, Turkey

In this study, technology-enhanced statistical problem-solving tasks designed by pre-service teachers (PTs) were examined. The PTs designed 28 tasks. The designed tasks were analyzed within the context of the Considerations for Design and Implementation of Statistics Tasks (C-DIST) components. It was revealed that the tasks were mostly designed within the framework of the learning goal of "statistical questions-making interpretations based on the measures that serve to represent the data and the forms of representation" and that mostly real, multivariate and large data sets were used. In addition, it was observed that the context was employed in order to complete the prepared tasks and the tasks mostly included the entire investigation cycle. It was determined that the prepared tasks were mostly at Level B, followed by the tasks at Level A and Level C. In light of the results obtained, inferences were made for preparing PTs to teach statistics.

Keywords: task design, technology, statistical problem solving, pre-service teachers

ARTICLE DETAILS

LUMAT Special Issue
Vol 11 No 3 (2023), 19–48

Received 10 February 2023
Accepted 15 September 2023
Published 27 September 2023

Pages: 30
References: 59

Correspondence:
nadideylmz70@gmail.com

https://doi.org/10.31129/LUMAT.11.3.1936

1 Introduction

Individuals need to learn how to read and analyze data, because data are encountered in all areas of life and it is necessary to make decisions based on data (Bargagliotti et al., 2020; Boaler & Levitt, 2019; Wild et al., 2018). This emphasis has found reflections in curricula (e.g., Australian Curriculum, Assessment and Reporting Authority [ACARA], 2015; Ministry of National Education [MoNE] (2018) and statistics has taken its place as a learning area in the curricula. Institutions giving direction to statistics teaching and curricula emphasize that students should experience statistical problem solving (SPS) in the teaching process (Bargagliotti, et al., 2020; Franklin, et al., 2005; MoNE, 2018). Undoubtedly, the tasks in the teaching process play a key role in helping students acquire these targeted statistical skills (da Ponte, 2011; Franklin, et al., 2015; Shaughnessy, 2007).

Students often encounter statistical tasks in textbooks in the teaching process (Braswell et al., 2005). However, it has been observed that the tasks in the textbooks mostly focus on the analysis of data rather than SPS (Balcı, 2023; Bargagliotti et al., 2020; Jones & Jacobbe, 2014; Jones et al., 2015) and take smaller datasets into their centre (Weiland, 2019). However, in real life, students need to work with large data

Published by the University of Helsinki, Finland / LUMA Centre Finland | CC BY 4.0

sets (Casey et al., 2021). At this point, the tasks prepared by teachers play a key role (Bakogianni, 2015; Dierdorp et al., 2011; Garfield, 1995; Shaughnessy, 2007). However, studies show that teachers have difficulties in designing statistical tasks (Bakogianni, 2015; Casey et al., 2020; 2021; Chick & Pierce, 2008; Rossman et al., 2006). In the current study, it is aimed to examine the technology-enhanced SPS task design assignments designed by PTs who will be the teachers of the future.

2 Literature Review

In this section, first, the features of high-quality SPS tasks are mentioned. Then, studies focusing on the features of statistical tasks designed by mathematics teachers/PTs are presented.

2.1 Features of high quality SPS tasks

High quality statistical tasks in teaching processes should be designed in such a way as to allow students to experience SPS (Bargagliotti et al., 2020; Burgess, 2011; Franklin et al., 2007). SPS consists of the following stages; formulating a statistical question, collecting or considering data, analyzing data and interpreting the results. (see. Figure 1).

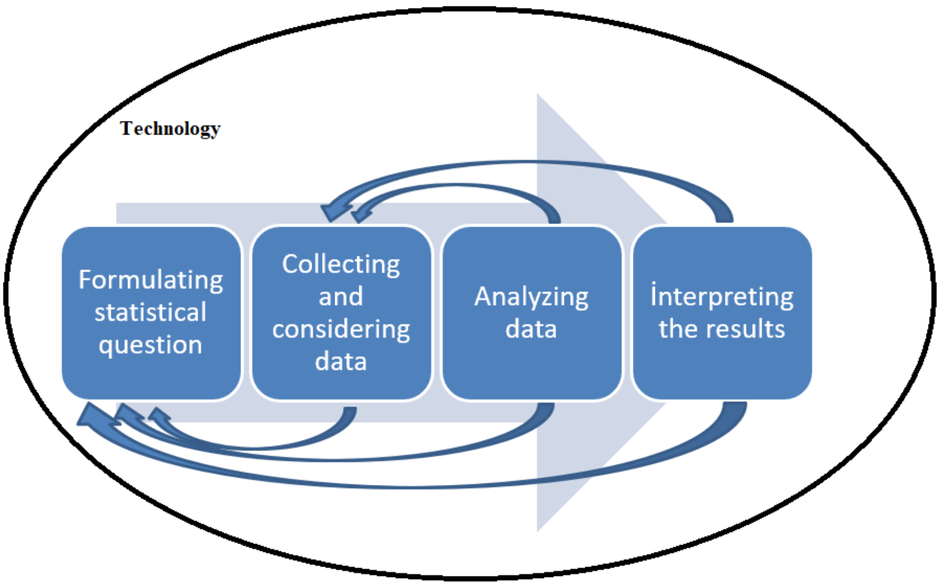


Figure 1. The relationship between technology and statistical problem solving (Adapted from Bargagliotti et al., 2020).

At the stage of formulating a statistical question the problem is defined clearly and the problem is shaped in a way that pays attention to variability and real-world context (Cobb & Moore, 1997; Franklin et al., 2005; Scheaffer, 2006). Collecting or considering data includes planning for the data to be collected or considered and implementing this plan (Franklin et al., 2005). Another point to be considered is that the data sets collected or selected should reflect the nature of daily life (Bargagliotti et al., 2020; Lee, 2019). At the third stage, analyzing data, the analysis of the data takes place by choosing the appropriate methods (Franklin et al., 2005). As at every stage, context should be taken into account when deciding on the appropriate data analysis method (delMas, 2004). Representing the data with different means of representation - in other words, transnumeration - or creating different meanings by interpreting a representation according to different perspectives takes place in this process (F. Curcio, 1987; F. R. Curcio, 1989; Shaughnessy, 2007; Wild & Pfannkuch, 1999). At the last stage, interpreting the results, the results are interpreted by considering variability and context and these results are associated with the initial research question (Bargagliotti et al., 2020; Franklin et al., 2005). While reading the data can reveal the information appearing on the graph, in reading between [or within] the data, the relationships between different components or data points can be defined. If inferences or generalizations are to be made based on data, reading beyond the data is required (F. Curcio, 1987; Friel et al., 1997). Reading behind the data can be performed to reveal contextual explanations about the trend in the data (Shaughnessy, 2007; Shaughnessy et al., 1996). In addition, attention is drawn to the use of technological software in order to make the SPS more effective (Bargagliotti et al., 2020) because technological tools allow creating graphical representations and producing numerical summaries of data and simulations. This paves the way for conceptual understandings to be placed in the centre by focusing more on statistical concepts and data (Bargagliotti et al., 2020; Franklin et al., 2005; 2015). Researchers working in this field have prepared a framework that reveals how the tasks to be used and implemented in the teaching process should include the SPS (Tran & Lee, 2015).

2.2 Features of the statistical tasks designed by mathematics teachers/PTs

Tasks plays a decisive role in the conduct of statistics teaching as targeted (e.g., Dierdorp et al., 2011; Shaughnessy, 2007). Attention is drawn to the need for PTs to prepare high-quality statistical tasks in mathematics teacher education (da Ponte, 2011; Franklin et al., 2015). Differences between mathematics and statistics, the role of context in statistics and the differences in the process of interpreting data make the preparation and implementation of statistical tasks even more critical (Bakogianni, 2015; Rossman et al., 2006). However, it has been revealed that there are limited studies on how teachers/PTs prepare statistical tasks, and in these limited studies, it has also been revealed that teachers/PTs have difficulties in designing statistical tasks (e.g., the role of context) (Bakogianni, 2015; Casey et al., 2020; 2021; Chick & Pierce, 2008; Rossman et al., 2006). Chick and Pierce (2008) gave data sets to PTs and asked them to create statistical questions and a hypothetical lesson plan for sixth graders using these data sets. It was observed that the majority of the PTs (81%) created questions asking to simply read or interpret the information shown by a table or graph, while less than half (41%) of the questions they prepared were aimed at identifying information that is not immediately visible from the data or making inferences from the data. More than half of the lesson plans prepared by the PTs directed the students to the data set. However, few of the lesson plans (23%) included the continuous and effective use of the data set to reveal statistical concepts, while the other lesson plans did not include the data set in a meaningful way. Casey et al. (2020) examined the tasks prepared by PTs and concluded that most of the tasks prepared by PTs contained large, multivariate, real datasets and allowed making associations with the context. In addition, the attempts of PTs to structure the tasks in a way that allows for a SPS have attracted attention. Casey et al. (2021) focused on the strengths of the tasks designed by PTs and the aspects that need improvement. They stated that the strengths of the tasks designed by PTs included the use of large and multivariate data sets, constant connection with the context and students' involvement in many parts of the SPS and their use of multiple data presentations. The aspects of the tasks that need improvement were expressed as adopting a mathematical approach instead of a statistical approach, focusing on ambiguous questions or numerical calculations as well as issues related to statistical content. Bakogianni (2015) focused on the stages of mathematics teachers preparing, implementing and reflecting on statistical inquiry tasks. Statistical context has proven to be not only an elusive learning goal, but also a significant teaching challenge. It has been determined that teachers' familiarity with the content and teaching of statistics, students' prior statistical knowledge, classroom reality problems and the stochastic context of statistical problems affect the preparation and implementation of the tasks and reflection on them. Collaboration and interaction among teachers provided the opportunity for teachers to gain a deeper understanding of statistical concepts and procedures, to identify specific learning objectives related to statistical content, and to be aware of learning difficulties associated with them. The results show that PTs have various difficulties in designing statistical tasks.

3 Rationale of the study

Targeted statistics education "depends to a large extent on the teachers who will bring them to life in the classroom" (Franklin et al., 2015 p.1). It is necessary for teachers to have many skills, such as being able to plan, conduct and evaluate the teaching of statistical concepts in the classroom environment (Franklin et al., 2007; 2015, Groth, 2007; 2013). One of the important factors determining the effectiveness of students' learning is the tasks used (Carver et al., 2016). It can be said that teachers generally tend to use the tasks from various sources (e.g., textbooks, online) (Casey et al., 2020; Shapiro et al., 2019). However, it has been revealed that some tasks in these resources focus on the calculation of concepts and analyzing the data and include small data sets, instead of dealing with SPS in its entirety (Balcı, 2023; Jones & Jacobbe, 2014; Jones et al., 2015; Weiland, 2019). Another emphasis is that statistical data are intertwined with technology and that the tasks used and the SPS should be integrated with technology when feasible (Bargagliotti et al., 2020). Many studies that draw attention to the fact that technology is one of the important components that affect the quality of statistics education also support this (Garfield, 1995; Lee et al, 2014; Neto, 2017; Suhermi & Widjajanti, 2020; Tishkovskaya & Lancaster, 2012). Seen from this perspective, it becomes important for teachers to have the necessary knowledge and skills on how to create high-quality tasks (Casey et al., 2020; 2021). Researchers pointing out that the reasons for the difficulties experienced by teachers / PTs should be examined in depth state that the lack of knowledge of PTs about SPS has the potential to affect the structure of the tasks they prepare (Bakogianni, 2015; Casey et al., 2020; 2021; Chick & Pierce, 2008). PTs' design statistical question (Burgess, 2007; Leavy & Frischemeier, 2022), data collection (Hannigan et al., 2013; Lovett & Lee, 2018), data representation or interpretation (Casey & Wasserman, 2015; Hannigan et al., 2013) may also affect the prepared tasks. In other words, if PTs have difficulties in carrying out the SPS, it is likely that this will affect the statistical tasks they prepare (Casey et al., 2020; 2021). Seen from this perspective, the examination of the statistical tasks prepared by PTs has the potential to provide mathematics educators with important information about PTs (Casey et al., 2020; 2021; Chick & Beswick, 2018). With the current study, it is thought that it will be revealed which points PTs can easily deal with while preparing statistical tasks and at which points they need improvement. On the basis of the results obtained, mathematics educators can better organize the content of undergraduate courses and more effectively conduct these courses (e.g., teaching statistics and probability, statistics). In addition, another point that should be taken into consideration is that limited research has been conducted in this field. The existing research has largely focused on the development of the content knowledge of PTs (Peck et al, 2013; Perkowski & Perkowski, 2007). In this context, it is thought that the results obtained will contribute to the literature. Thus, an answer to the following research question was sought: How are the technology-enhanced SPS task design assignments designed by PTs?

4 Method

4.1 Research design and participants

Since the purpose of the current study was to examine the structure of the statistical task design assignments prepared by PTs within the scope of a teacher education program, the case study design was employed. Case study allows obtaining and examining in-depth information about the case of interest in line with the research problem (Merriam, 2009; Putney, 2010). In the current study, it was aimed to examine the task design assignments prepared by PTs. To this end, the unit of analysis of the study was determined as 28 task design assignments prepared by 56 PTs participating in the study in groups of two.

4.2 Context of the study and data collection

This study was carried out in the Department of Mathematics at a state university in a city located in the Central Anatolian region of Turkey. PTs who graduate from this department can work as a mathematics teacher at the middle school level (11-14 years old) of public or private institutions. The program is a four-year program and the language of instruction is Turkish. In the first two years of the four-year program, PTs mainly take content knowledge courses (e.g., Analysis, Algebra) and in the last two years, they mainly take pedagogical content knowledge courses (e.g., Teaching practice, Teaching numbers). This study was conducted within the scope of the "Probability and Statistics Teaching" course, which is a compulsory course to be taken in the sixth term of the mathematics teaching program. The course mainly focused on the teaching of probability and statistics concepts and aimed to improve PTs' knowledge and skills about these concepts and how this knowledge could be reflected in the classroom environment. The course also placed the SPS into the centre of the subjects taught (Bargagliotti et al., 2020). In addition, approaches to teaching statistics and how to implement an effective statistics teaching (Ben-Zvi, 2011; Cobb & McClain, 2004) were discussed with the PTs. In the following weeks, each stage of SPS (formulating a statistical question, collecting or considering data, analyzing data and interpreting the results) was handled. In addition, the technological software that supports the teaching of statistics and how this software could support the teaching process were discussed. Common Online Data Analysis Platform (CODAP) (http://codap.concord.org), web-based educational software, was preferred because it is free and accessible, and students were informed about how to use this software. Furthermore, sample tasks were examined (Concord Consortium, 2019). In the 10th week of the lesson, the PTs were asked to design a statistical task design assignment using CODAP in such a way as to develop students' statistical thinking as a group. This task design assignment was requested to consist of two main parts (Figure 2).

1) Introductory Information

Brief Summary of Task:

Learning goal of the task you have prepared:

Target audience:

Time required:

Materials:

CODAP Link: (Add a link to your data and analyses)

2) Component of a Statistics Task (While structuring your task, suppose that students are solving it)

Posing Statistical questions (e.g., What is your question? What did you pay attention to while posing?)

Data Collection (e.g., How did you obtain the data?, From whom did you collect the data?)

Data Analysis (e.g., How did you analyze the data you collected/used? What representations, measurements (e.g., central tendency, dispersion) did you use?

Interpreting the results (e.g., What do the results tell us?)

Figure 2. Task design assignment

These parts are defined as (1) task summary, target audience, required time, materials, learning goal, CODAP link and (2) Component of a Statistics Task (data, context, SPS). The PTs were left free to use whichever dataset they wanted to use while designing their tasks. It was stated that they could create the data themselves if they wished, or they could use ready-made data sets. The PTs were also said that they could benefit from various websites (e.g., Turkish Statistical Institute (TUIK), Sample CODAP Datasets). In addition, it was explained to the PTs that they could get support from various sources (e.g., curriculum, academic resources) while preparing task design assignments. The PTs designed a total of 28 tasks.

4.3 Data analysis

The task design assignments developed by the PTs were analyzed in the context of C-DIST components developed by Tran and Lee (2015, p.1–2). The PTs focused on preparing the tasks, they did not engage in task implementation. Therefore, the task design assignments prepared by focusing on the components in the framework of "Considerations for Written Task" were analyzed. These components are shown in Table 1.

Table 1. Considerations for written task framework (Tran & Lee, 2015 p.1–2)

Component of a Statistics Task	Questions to Consider
Learning Goal	What learning goals does the task aim for students to accomplish? Does the task focus on answering questions that are statistical or mathematical? e.g., Does the task ask students to use computations or graphs? Are these in support of analyzing data to make a decision? or is the use of an algorithm or creation of a graph the focus?
Data	Does the task call for the use of data (either to collect or use already collected data to answer)? Does the data appear to come from a real source?
Context	Is context a salient part when solving the problem? Is the context likely to be of interest to the students engaging in the task?
SPS	Does the task address only one phase of a SPS, some phases, or all phases of the cycle? Consider the appropriate phases below as applicable to the intent of the task:
Pose	Is the question already posed (by teachers, or curriculum developers) or do students have opportunities to pose statistical questions based on their interest? What type of variability does the task attend to?
Collect	Does the task offer opportunities for students to plan to collect data: sampling, sample size, attribute, and measurement? Do students conduct the data collection? Does the task provide a context so that students are aware of the measurement issues and how data were collected?
Analyze	

Interpret	Does the task offer opportunities for students to decide on the types of graphical representation and or numerical statistics to use when analyzing data? Does the task afford students to use alternative representations to shed light on the trends of data? Does the task ask students to incorporate context when making claims/inferences about the data? Does the task expect students' claims to account for uncertainty?

Each component in Table 1 was detailed within itself and criteria were created. First, two main criteria were determined for the first component, the learning goal. It was questioned which learning objectives the prepared task aimed to make students accomplish. If the learning goal focused only on graphing (e.g., represent this data with a line graph) or just doing calculations (e.g., what is the mode of this data?), it was coded as a mathematical question. Learning goals where a context was used and statistical situations were required to be evaluated were coded as statistical questions. Statistical questions, on the other hand, were evaluated within the framework of two subcomponents. If the question asked for interpretation by using the measures and types of representation that serve to represent the data, that question was coded as "Interpretation based on the measures and types of representation used to represent the data-statistical question". If the question asked for making inferences by using the measures and types of representation that serve to represent the data in the question, that question was coded as "Making inferences based on the measures and types of representation used to represent the data-statistical question".

The prepared tasks were analysed in terms of the characteristics of the data and the need for data in the context of the data component. If the data used were directly collected from a real source, they were coded as "Data come from real source-primary data". If the data were obtained from a real source (e.g., websites, OECD), they were coded as "Data come from real source-secondary data". If the data were fabricated by PTs, they were coded as "Data come from hypothetical". Another point examined was whether data were needed to complete the prepared task design assignments. In this context, the "Use of data to complete the task" code was created.

In terms of the context, what the context addressed contains was analyzed. In this connection, based on the contexts provided in the task design assignments prepared by PTs, components such as "Contexts related to the students themselves", "science", "health", "social" and "education" were created and the task design assignments were examined within this framework. In order to determine whether the contexts in the task design assignments were contexts that could capture the attention of students, contexts found in the textbooks and reference books commonly used by students were taken into account. In addition, the contextual information contained in the contexts was also examined in order to reveal in more detail how the tasks are related to the data's context. Contextual title, contextual attribute names, multivariate nature of datasets were discussed in this regard. In addition to this, task design assignments were analyzed according to including/not including information about the source of the data. In other words, it was analyzed whether the source of the data was included in task design assignments. Each task design assignment prepared by the PTs was evaluated according to these criteria. While evaluating the tasks in terms of SPS components (pose, collect, analyze, interpret), the framework created by Bargagliotti et al. (2020, p.16–19) was adopted. This framework described each component at Level A, B and C. The analysis also included the examination of the shared CODAP link associated with the task design assignments. In this way, the types of data representations involved in the link and whether information on how to use CODAP or the data source was included were determined. After the data were coded, another researcher was asked to code the data independently. After the two codings, the researchers came together and analyzed the task design assignments in the context of the components and the points of disagreement were discussed until a consensus was reached.

5 Findings

5.1 Learning goal

The task design assignments prepared by the PTs were evaluated in terms of learning goals and Table 2 was created.

Table 2. Learning goals of the tasks

Learning goal	Frequency	Percentage
Mathematical calculation/graph construction-mathematical question	1	4%
Interpretation based on the measures and types of representation used to represent the data-statistical question	23	82%
Making inferences based on the measures and types of representation used to represent the data-statistical question	4	14%

It was observed that the prepared task design assignments mainly included the learning goal of "statistical questions- Interpretation based on the measures and types of representation used to represent the data" (82%). For example, the learning goal in one of the tasks was expressed as “determining whether the use of the left or right hand is related to gender”. Only in one task, the learning goal of “mathematical questions-mathematical calculation/graph construction” was at the forefront. The learning goal of this task was determined as “calculating the average lifespan of mammals”. In four of the tasks prepared by the PTs, the learning goal was “Making inferences based on the measures and types of representation used to represent the data-statistical question”. For example, the learning goal in one of these tasks was stated as “estimating the foot length according to the height of individuals”.

5.2 Data

The task design assignments prepared by the PTs were evaluated in terms of learning goals and Table 3 was created.

Table 3. Features of the data in the tasks

Data	**Frequency**	**Percentage**
Data come from real source-primary data	13	46%
Data come from real source-secondary data	14	50%
Data come from hypothetical	1	4%
Need for data		
Use of data to complete the task	28	100%

In half of the task design assignments prepared by the PTs, data collected from real secondary sources were preferred. These data sets were obtained from TUIK data, sample data sets in CODAP, and various internet sites. In 46% of the tasks, the data consisted of primary sources collected by the PTs themselves. In only one task, it was determined that the PTs created the data themselves, that is, they obtained hypothetical data. In addition, it was observed that all the designed tasks required the use of data. Then, the number of data used in the tasks was analyzed. In the tasks, while the minimum number of cases was 27 the maximum number of cases was 900. The mean of the number of cases was 274, and the median was 218. Based on these results, it can be said that they preferred to use large data sets in most of the tasks. It was also noted that both categorical and quantitative (numerical) variables were included in task design assignments.

5.3 Context

The contexts of the task design assignments prepared by the PTs was evaluated in terms of their being interesting/suitable for the level of the students and the consideration of the context in the process of solving the problem and Table 4 and Table 5 was created.

Table 4. Features of the context in the tasks

Data	Frequency	Percentage
Contexts related to the students themselves (height, favourite subjects, number of siblings..)	15	54%
Science (atomic radius, mammals...)	5	18%
Health (Covid 19, heart attack..)	4	14%
Social (seasonal workers, population, tobacco use...)	3	11%
Education (budget in education…)	1	3%

It was determined that all the task design assignments prepared by the PTs were interesting and could attract the attention of students. About half (54%) of the tasks prepared by the PTs were found to include contexts related to the students themselves. This was followed by science (18%) and health (14%) contexts. Social context (11%) was preferred in one of every ten tasks prepared. The least preferred context was found to be education in the task design assignments prepared by PTs (3%).

Another point analyzed was the need for context in order to solve the statistical question. In this connection, Table 5 was created.

Table 5. Using the context to solve the question in the tasks

Data	Frequency	Percentage
Including contextual information (e.g. contextual title, contextual attribute names)	28	100%

When evaluated in terms of the sub-components showing that context was used to solve the statistical question, it was remarkable that all of the task design assignments included contextual information. An example of this is the title "the number of seasonal workers in the provinces of..." in a graph created for a statistical question focusing on how the number of seasonal workers changes across different provinces. In a table containing the data in another task, the 10-letter figure refers to the sum of the letters in the name and surname of one of the students. When evaluated in terms of multivariate nature of datasets, it was revealed that there are approximately 4 (mean

4.1) attributes per case. The obtained results allow the interpretation that multivariate data sets were used in the tasks. When the designed tasks were evaluated in terms of the source of data, Table 6 was presented.

Table 6. Using the context to solve the question in the tasks

Data	Frequency	Percentage
Including information about the source of the data	24	86%
Not including information about the source of the data	4	14%

Table 6 was revealed that explanations were made for the data obtained from the primary sources (for example, we wanted to measure the foot length and height of the university students who wanted to be at the university for a week), and in the data sets obtained from the secondary sources, the source was included, except for 4 tasks. On the other hand, in the task constructed from the hypothetical data, it was stated that they created these data themselves, since they could not be accessed from any source.

5.4 SPS

The task design assignments prepared by the PTs were evaluated in terms of SPS and Table 7 was created.

Table 7. Using SPS in the tasks

SPS	Frequency	Percentage
Including one or more stages of SPS	3	11%
Including the whole SPS	25	89%
SPS levels		
Level A	8	28%
Level B	16	58%
Level C	4	14%

While it was observed that the majority of the tasks (89%) prepared by the PTs included the whole SPS, 11% did not include the stage of interpreting the results. The majority of the tasks prepared by the PTs were found to be at Level B (58%), followed by Level A (28%). Only four of the tasks prepared by the PTs were found to be at Level C.

Level A tasks

Although statistical questions related to the tasks prepared at this level included small groups (e.g., PTs in a classroom, students in a classroom, some cities), it was observed that these groups were well defined. Some sample questions in these tasks are "What kind of sports do the students in our class like?, What is the distribution of the colours of the sweaters worn by the PTs?, Did the Black plague or Covid-19 cause more deaths in the cities of Moscow, Venice, London, Beijing, Paris and Warsaw? What is the average lifespan of mammals?". In addition, it was noted that the contexts used were chosen in a way that would attract the attention of students.

It was observed that the data sources used to answer these questions in the tasks were obtained from both primary and secondary sources. For example, the PTs collected data from their classrooms to answer the research question about what kind of sports the students in our class like, while they used the CODAP sample dataset to answer the question about the average life span of mammals. They noted down the data they collected/used in excel files or papers. It was observed that the PTs used both categorical (e.g., sports) and numerical (e.g., weight) variables. In addition, it was seen that they were aware of how the variable (e.g., the sum of the number of letters in the names) in the tasks they prepared was distributed, that is, how many times a certain result occurred.

It was observed that the PTs used different representations such as tables, bar charts, picture graphs, and dotplots while analyzing the data they collected in order to answer the research question they prepared, as well as taking into account the measures of central tendency (e.g., arithmetic mean).

It was seen that the PTs interpreted the data they analyzed by taking into account the group they dealt with. For example, "In the city of Moscow, the number of deaths caused by Covid 19 is higher than the number of deaths caused by Black Plague. However, the number of deaths caused by Black Plague in Beijing city is higher than the number of deaths caused by Covid 19". The mean was calculated in a task but it was observed that no comment was made on what this mean meant. Below are presented two sample tasks for this level.

In the first task, the PTs created the following research question to be answered on the basis of their own classroom "A sports tournament is planned to be held at the end of the year for PTs. For this purpose, it should be determined which sport is liked by the PTs. What kind of sports do the PTs in our class like?". To collect data for this purpose, they created the following survey question "What is your favourite sport?". The questionnaire prepared to collect data is given below in Figure 3.

Gender Male ☐ Female ☐

Age

What is your favourite sport?

Basketball ☐

Football ☐

Tennis ☐

Volleyball ☐

Figure 3. Data collection tool in the task

Here, it is seen that the data in the data collection tool are categorical data. It can be said that since the survey question asked to choose between four options, it became easier to organize and analyze the data. However, a limitation can be pointed out here. The fact that sports branches are limited to four options will make it difficult for PTs who do not like one of these sports and like another sport (e.g., athletics) to choose. Below in Figure 4 are given the raw data collected by the PTs.

cases (75 cases)									
indeks	Ogrenciler	Favourite sports type	gender	age	indeks	Ogrenciler	Favourite sports type	gender	age
1	Zuleyha	Volleyball	Female	22	29	Naile	Basketball	Female	21
2	Hasan	Football	Male	21	30	Turkan	Football	Female	21
3	Berkay	Football	Male	21	31	Buse	Basketball	Female	21
4	Ahmet	Tennis	Male	20	32	Nur	Basketball	Female	21
5	Burak	Basketball	Male	21	33	Tugba	Tennis	Female	21
6	Meryem	Basketball	Female	21	34	Oguzhan	Football	Male	21
7	Burcu	Volleyball	Female	20	35	Zeliha	Tennis	Female	21
8	Hasibe	Tennis	Female	20	36	Muhammed	Football	Male	21
9	Yasin	Football	Male	20	37	Ali	Volleyball	Male	21
10	Emir	Basketball	Male	21	38	Yagmur	Tennis	Female	21
11	serife	Volleyball	Female	21	39	Merve	Tennis	Female	21
12	Elif	Tennis	Female	21	40	Necati	Football	Male	20
13	Fatma	Volleyball	Female	21	41	Yusuf	Football	Male	20
14	Hilal	Football	Female	21	42	ilayda	Tennis	Female	20
15	Beyza	Volleyball	Female	21	43	Melike	Basketball	Female	20
16	Abdulbaki	Basketball	Male	21	44	Ahmet	Football	Male	20
17	Emine	Basketball	Female	20	45	Selim	Football	Male	20
18	Duygu	Volleyball	Female	21	46	ikranur	Volleyball	Female	21
19	Ayse	Basketball	Female	20	47	Humeyra	Volleyball	Female	21
20	Naz	Tennis	Female	20	48	Burak	Football	Male	20
21	Fatma	Tennis	Female	21	49	Kemal	Basketball	Male	20
22	Rahime	Basketball	Female	21	50	Semanur	Football	Female	21
23	Tarık	Basketball	Male	21	51	serife	Tennis	Female	21
24	Elif	Tennis	Female	20	52	Asli	Tennis	Female	21
25	Zehra	Basketball	Female	20	53	Cennet	Volleyball	Female	20
26	Hayriye	Volleyball	Female	20	54	Fadime	Tennis	Female	20
27	Sila	Tennis	Female	22	55	Beyza	Basketball	Female	20
28	Hamide	Volleyball	Female	21	56	Fatmanur	Volleyball	Female	20

Figure 4. Raw data collected for the task

Based on these collected raw data, the PTs created horizontal and vertical dotplots to represent the data and they also represented the data with a table.

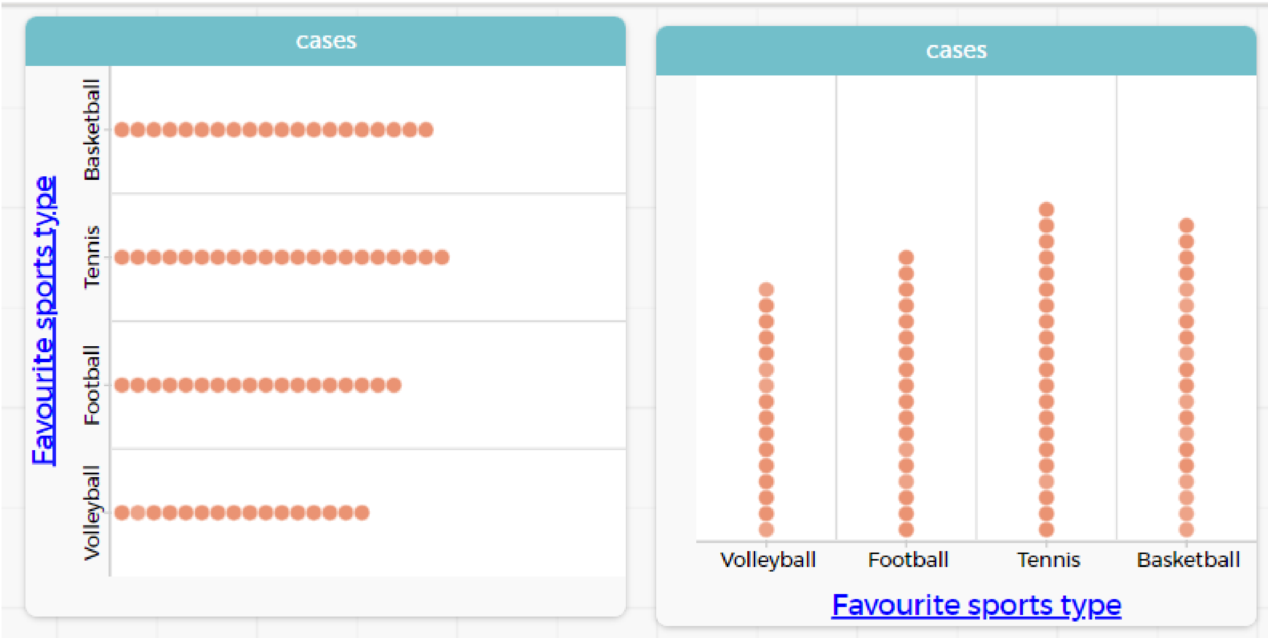


Figure 5. Types of representations used in the task

The PTs interpreted the data they analysed in the tasks they prepared as follows; “The numbers of the sports branches that the PTs like are very close to each other, the PTs preferred tennis the most, we can say that when one of the PTs is chosen randomly, his/her probability of liking volleyball is less than his/her probability of liking tennis.” These interpretations can be considered as an indicator of what they are doing is reading the data and reading between the data.

Another sample task prepared by the PTs at Level A was about the lifespan of mammals. In the task they prepared, the PTs asked the question “What is the average lifespan of mammals?”. The data set used was a sample data set in CODAP and retrieved from https://codap.concord.org/app/static/dg/tr/cert/index.html.

Memeliler

Mammals (27 cases)

index	Memeli	Order	Yaşam…i (years)	Uzunluk…(meters)	Ağırlık (kg)
1	Afrika Fili	Probosc…	70	4	6400
2	Asya fili	Probosc…	70	3	5000
3	Yarasa	Chiropt…	19	0.1	0,02
4	Yunus	Cetacea	25	3.5	635
5	Çita	Carnivora	14	1.5	50
6	Şempa…	Primate	40	1.5	68
7	Kedi	Carnivora	16	0.8	4,5
8	Eşek	Perisso…	40	1.2	187
9	Zürafa	Artioda…	25	5	1100
10	Kurt	Carnivora	16	1.6	80
11	Fok Balı…	Pinnipe…	30	2.1	275
12	Sincap	Rodentia	9	0.3	0,1
13	At	Perisso…	25	1.5	521
14	Hamster	Rodentia	3	0.1	0,03
15	Keçi	Primate	80	1.9	80
16	Jaguar	Carnivora	20	1.8	115
17	Katil Ba…	Cetacea	50	6.5	4000
18	Aslan	Carnivora	15	2.5	250
19	Keseli sı…	Didelph…	5	0.5	5
20	Armadil…	Xenarth…	10	0.6	7
21	Baykuş …	Primate	12	0.4	1
22	Patas M…	Primate	20	0.9	13
23	Domuz	Artioda…	10	1	192
24	Antilop	Artioda	10	0.9	70

Memeli Örnek Kılavuzu

Bu veri seti 27 memeli hakkında bilgi içermektedir. Onlar hakkında neler öğrenebileceğinizi görmek için tabloya ve grafiklere bakın.
Tabloda memelilerin yaşam sürelerini, boy uzunluklarını, ağırlıklarını, yaşam alanlarını ve beslenme türlerini görebilirsiniz. Sadi Bey yeni bir yatırım için bir arsa almıştır ve bu arsaya yeni bir hayvanat bahçesi kurmak istemektedir. Bunun için her hayvan türünde uzman zoologlar ile çalışmaktadır.

Hayvanat bahçesine getirtilecek olan memeli hayvanlar içinse Zoolog Kamil Bey ile çalışacaktır. Kamil Bey'den belli standartlara göre hayvanları listelemesini istemiştir. Kamil Bey'in vereceği liste doğrultusunda taşıma konusunda en az maliyet oluşturacak hayvanları seçecektir.

Kamil Bey'in en çok dikkat ettiği özellikler hayvanın boy uzunluğu, ağırlığı ve yaşam süresidir. Sizce memelilerin ortalama yaşam süresi nedir?

Figure 6. Raw data and related context for the task

The PTs calculated the average lifespan of mammals here. They found the average as 24.85. However, they did not make any interpretations on this result. Here, the PTs were expected to make evaluations about which mammal lifespan is closer to the average and which mammal lifespan is farther from the average because such interpretations are also evaluations of variability across the data set obtained.

Level B tasks

In the tasks prepared at this level, it was observed that statistical questions were formed for comparison (e.g., do the types of music that students like differ between classes?) and association (e.g., is using right or left hand related to gender?) between variables based on a larger sample. It was also seen that they included questions that required examining the change of a variable over time (for example, how the average life expectancy of women in Turkey changed by years). In addition, PTs designed questions that would include two categorical (do students who like bananas tend to like/dislike strawberries?), two numerical (Is the height of students related to their jumping height?) variables and one categorical and one numerical variable (Is using a computer program effective on students' statistical exam grades?). In addition to collecting first-hand data, they created research questions using data sets obtained from online sources and websites (e.g., TUIK). It was noted that the collected/obtained data were recorded in the Excel program. They also made arrangements to perform random assignments to control certain traits. For example, a research question "Does going to the support course affect the number of words that students read in 1 minute?" was asked. In this context, students studying in ….primary school 1/A, 1/B, 1/C and 1/D classes were determined. Although the students in these classes were not randomly selected, the students were randomly assigned while conducting the relevant experiment.

It was revealed that various representations (e.g., bar graph, scatterplots, two-way graph, dot plot, mosaic plot, a time series plot) were used to analyze the collected data, measurements (e.g., mean absolute deviation (MAD), measurements of central tendency) were used to describe the distribution and measurements (e.g., correlation coefficient) were used to elicit the relationship between two variables.

It was seen that while interpreting the analyzed data, they used expressions to look at reading beyond the data as well as reading the data and reading between the data. It was also observed that they made comments to compare the results for different conditions in an experiment (e.g. how using/not using a computer program affects statistical grades). In addition, it was observed that they stated that although the selected samples were larger than the samples at Level A, they still might not represent the population (for example, although we evaluate the average life expectancy of women in Turkey, we cannot say that the average life expectancy of women all over the world is like this). A sample task is presented below.

In the sample task, the PTs focused on the periodic table of the elements. Based on the various properties of the elements in the periodic table, they prepared research questions to reveal the relationship between these properties. For example, they posed the following research question; "Is there a relationship between the melting points of the elements and their boiling points?" To this end, they obtained the data from a secondary data source. The data were obtained from https://codap.concord.org/app/static/dg/tr/cert/index.html. An example of the data they handled is given below in Figure 7.

List of Periodic Table Elements

Periodic Table Elements (112 cases)

AtomicNumber	Name	aAtom Ağırlığı (amu)	attr	Erime Noktası (˚ C)	Kaynama Noktası (˚ C)	Density (g/cc)	Year Discovered	İyonlaşma Enerjisi (eV)	Table Row	Table Column	ChemicalSeries	Atom yarçapı
1	Hydrogen	1.01	H	-259.14	-252.9	0	1766	13.6	1	1	Nonmetal	
2	Helium	4	He	-272	-268.6	0	1895	24.59	1	18	Noble gas	
3	Lithium	6.94	Li	180.54	180	0.53	1817	5.39	2	1	Alkali metal	
4	Beryllium	9.01	Be	1278	2970	1.85	1798	9.32	2	2	Alkaline earth ...	
5	Boron	10.81	B	2300	2550	2.34	1808	8.3	2	13	Metalloid	
6	Carbon	12.01	C	3500	4827	2.25		11.26	2	14	Nonmetal	
7	Nitrogen	14.01	N	-209.9	-196	0	1772	14.53	2	15	Nonmetal	
8	Oxygen	16	O	-218.4	-1.4	0	1774	13.62	2	16	Nonmetal	
9	Fluorine	19	F	-219.62	-188	0	1886	17.42	2	17	Halogen	
10	Neon	20.18	Ne	-248.6	-249	0	1898	21.56	2	18	Noble gas	
11	Sodium	22.99	Na	97.8	883	0.97	1807	5.14	3	1	Alkali metal	
12	Magnesium	24.31	Mg	638.8	1090	1.74	1755	7.65	3	2	Alkaline earth ...	
13	Aluminum	26.98	Al	660.37	2467	2.7	1825	5.99	3	13	Poor metal	
14	Silicon	28.09	Si	1410	2355	2.34	1823	8.15	3	14	Metalloid	
15	Phosporus	30.97	P	44.1	280	1.82	1669	10.49	3	15	Nonmetal	
16	Sulfur	32.07	S	112.8	445	2.07		10.36	3	16	Nonmetal	
17	Chlorine	35.45	Cl	-100.98	-35	0	1774	12.97	3	17	Halogen	
18	Argon	39.95	Ar	-189.3	-186	0	1894	15.76	3	18	Noble gas	
19	Potassium	39.1	K	63.65	774	0.86	1807	4.34	4	1	Alkali metal	
20	Calcium	40.08	Ca	839	1484	1.55	1808	6.11	4	2	Alkaline earth ...	
21	Scandium	44.96	Sc	1539	2832	2.99	1879	6.56	4	3	Transition metal	
22	Titanium	47.88	Ti	1660	3287	4.5	1791	6.83	4	4	Transition metal	
23	Vanadium	50.94	V	1690	3380	6.1	1830	6.75	4	5	Transition metal	
24	Chromium	52	Cr	1857	2672	7.1	1797	6.77	4	6	Transition metal	
25	Manganese	54.94	Mn	1245	1962	7.4	1774	7.43	4	7	Transition metal	

Figure 7. The data set obtained for the task

They created scatterplots to analyze the data they obtained (See Figure 8).

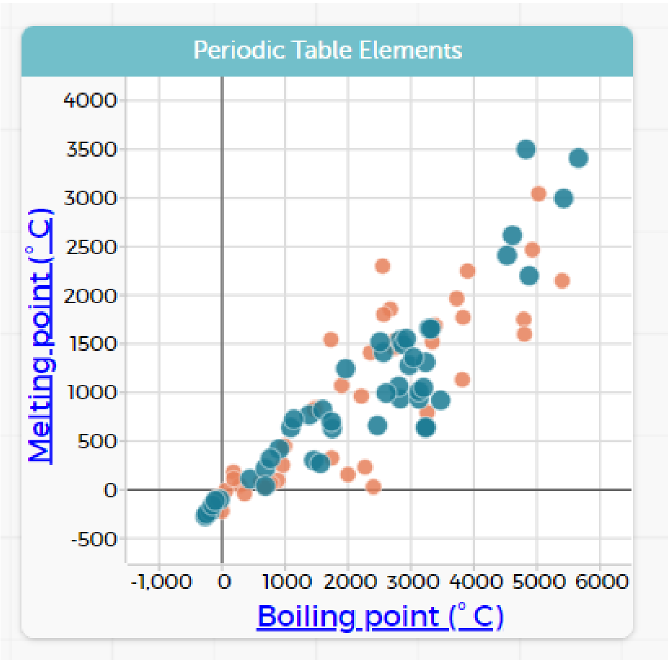


Figure 8. The scatterplot used for the task

Based on this graph, it was noted that the relationship between the melting point and boiling point of the elements was close to linear, that is, they made interpretations that the melting point of the relevant element would increase as the boiling point of the related element increased.

Level C tasks

It was noted that the statistical questions asked by the PTs in the activities they prepared at Level C included two or more variables and focused on causality and prediction. For example, the following questions can be given as examples to the questions at this level; “What is the relationship between the time of having a heart attack and its being fatal?, Is there a relationship between the height of individuals and the length of their feet? Can height be estimated from the foot length?, How has the working population in Turkey changed over the years?”. They preferred to use primary or secondary data sets to answer these questions. They were able to determine the appropriate method (e.g., survey research, observational studies and experiments) according to the research question. They became aware of the role of random selection when selecting samples and the fact that the random assignment in experimental assignments influenced cause-and-effect interpretations.

It was seen that they used high-level statistical concepts (e.g., population proportion (p), Pearson’s correlation coefficient (r), Quadrant Count Ration (QCR)) and high-level analysis methods (e.g., Chi-squared tests) when analyzing the data. In addition to making interpretation on what the estimation of a variable’s property means, they also used expressions to understand how the variables affect each other. Here, it can be said that PTs could make advanced interpretations (reading between, beyond and behind the data). Below is given a sample task.

In the sample task, the PTs formulated a statistical question; “Is there a relationship between the height of the individuals and the length of their feet? Can height be estimated from foot length?” In order to answer this question, PTs preferred to collect data themselves and prepared a questionnaire given below in Figure 9.

Gender Female ☐ Male ☐

Faculty :

Grade level :

Height : (cm)

Lenght of the feet : (cm)

Figure 9. Data collection tool for the task

Here, the PTs were aware that they could not reach all students in every faculty. Thus, they asked various eager students to fill out the questionnaire. They were also able to define and record the obtained variables.

As an example, regarding the problem of estimating the height from the foot length, the PTs measured and recorded the height and foot length of 100 university students randomly selected on the campus.

Person	Gender	Height(cm	Foot length(cm)	Person	Gender	Height(cm	Foot length(
1	K	160	25	51	E	175	26,8
2	E	173	26,7	52	K	168	25,6
3	K	162	22,7	53	K	169	26,7
4	E	173	27,3	54	E	176	27,6
5	K	158	24,2	55	E	173	27
6	K	160	23	56	K	163	24,6
7	K	168	25,3	57	E	176	24,6
8	E	178	27,8	58	E	172	24,6
9	E	172	26,4	59	K	171	25
10	E	168	24	60	K	165	22,6
11	E	170	26,9	61	K	166	24
12	K	167	26,5	62	E	178	24,9
13	K	162	24,2	63	E	177	27
14	E	175	27,2	64	E	176	26,8
15	E	173	26,8	65	K	151	23,6
16	K	162	22,7	66	E	184	28,1
17	E	174	27	67	K	159	24,1
18	K	163	24,3	68	E	165	24,3
19	K	157	23	69	E	169	24,6
20	E	157	24,6	70	E	171	26,9
21	E	175	24,8	71	K	181	26,8
22	K	169	24,8	72	E	177	25,6
23	K	158	24,5	73	E	178	25,7
24	E	175	28	74	E	177	25,7
25	E	172	26	75	K	164	24,3
26	K	160	23,6	76	E	176	24,9
27	E	168	24,4	77	E	170	24,6
28	E	176	25,1	78	E	166	24,3
29	K	162	24,7	79	K	166	27,2
30	E	170	27,5	80	E	162	22,5
31	K	163	24,9	81	E	167	22,6
32	E	185	27,9	82	K	164	25
33	E	173	28	83	E	178	28
34	K	167	24,2	84	E	177	25,2
35	E	183	29	85	K	172	25,2
36	K	160	22,7	86	K	163	23,6
37	K	163	22,5	87	E	167	23,9
38	E	186	28,4	88	E	176	27
39	K	166	22,6	89	K	150	22,8
40	E	173	25,3	90	E	174	27
41	E	176	25	91	E	178	29
42	K	160	22	92	K	172	23,7
43	K	160	23,4	93	E	167	25
44	E	176	24	94	K	159	22,8
45	K	154	23,3	95	E	167	23,5
46	E	176	27,3	96	E	169	24,6
47	E	173	27,1	97	K	156	24,3
48	E	162	25	98	K	167	22,6
49	K	165	25,3	99	E	178	24,7
50	K	160	23	100	K	165	25,6

Figure 10. Raw data collected for the task

They transferred the collected data to the CODAP program and calculated Quadrant Count Ratio (QCR) by drawing a scatterplot. They also included calculating the Pearson correlation coefficient (r), which takes into account the distance of the data from the mean lines. The results obtained showed a strong positive linear correlation between height and foot length. Based on this correlation, they wrote the equation for the least-squares line, with the help of technology to estimate the foot length from the height.

Foot length=0.1639*(height)-2.5

They evaluated whether the model prepared here was suitable and had a good fit and showed the positive and negative deviations in the data by drawing the fit line.

For example, the foot length of a student whose height is 185 cm

Foot length =0.1639*(185)-2.5=27.82

Figure 11. Residual plot and least-squares line

By analyzing the data, they argued that there was a linear correlation between foot length and height and that foot length could be used to predict height and they supported this argument with examples.

6 Discussion and conclusion

The current study aimed to examine SPS tasks prepared by the PTs using the CODAP dynamic statistics software tool. One of the motivations of the study was that many other studies (Casey et al., 2020; 2021; Langrall et al., 2017; Shaughnessy, 2007) drew attention to the difficulties experienced by PTs in this regard. In this connection, the PTs held various discussions on how to prepare SPS tasks within the scope of a course, and made examinations about the preparation of appropriate tasks and the selection of appropriate technological software. Then, the PTs prepared task design assignments. The prepared tasks were analyzed within the context the C-DIST framework.

When the task design assignments are evaluated in terms of learning goals, it can be said that the pre-service teachers tended to prepare statistical questions. From among the learning goals, interpretation based on the measures and types of representation used to represent the data was largely adopted. Making inferences based on the measures and types of representation used to represent the data was a less preferred learning goal by PTs in their task design assignments. Only one task focused on calculating the arithmetic mean, but this measure was carried out by taking the mathematical calculation into the centre. These findings can allow the interpretation that the PTs determined the goals of statistical tasks with a statistical approach. However, in a designed task, it was focused on calculating the arithmetic mean by adopting a mathematical approach. Casey et al. (2021) stated that contrary to the findings of the current study, the PTs mostly structured the tasks they prepared with a mathematical approach. When Chick and Pierce (2008) examined the lesson plans of the PTs, they observed that the majority of them made plans with the aim of simply reading or interpreting the information shown by a table or graph, and less than half of the questions they prepared were aimed at identifying information that was not immediately visible from the data or making inferences from the data.

When the data used in the tasks were examined, it was seen that in half of the task design assignments, the data came from the real source and the secondary data source was used, while in 46% of the tasks, the data came from the real source and were collected by the PTs themselves. The PTs' use of mostly real but secondary data can be evaluated under several headings. The first of these may be to provide information about various websites in the course so that PTs can access the data sets. The PTs, who gained knowledge about how to access the data sets, may have preferred secondary data sets for the tasks they prepared. It was observed that a hypothetical data set was preferred in only one task. The fact that the PTs could not reach the appropriate data set may have caused them to prefer the hypothetical data set. When the tasks were evaluated in terms of the number of cases, it can be said that the PTs preferred to use large data sets. In addition, it was observed that the data were multivariate and different attributes were taken into consideration. The use of large, multivariate and real data sets for effective statistics teaching is a point that has been emphasized in many studies (Bargagliotti et al., 2020; Casey et al., 2020; 2021; Franklin et al., 2015). In addition to studies that have reached similar results (Casey et al., 2020; 2021), there are studies that draw attention to different results (Chick & Pierce, 2008; Weiland, 2019). For example, Weiland (2019) revealed that small, univariate and bivariate tasks containing imaginary data are included in high school textbooks. In addition, another finding is that the task design assignments created by PTs were structured in such a way as to need data in order to be completed successfully. Parallel to this result, Chick and Pierce (2008) examined the lesson plans prepared by PTs in their study and stated that more than half of the lesson plans directed their students to the data set.

It was revealed that the PTs mostly tended to use the contexts related to themselves in the designed task design assignments, followed by the contexts in the field of science. It was observed that health and social contexts were preferred by PTs in approximately one out of every ten tasks. In the task design assignments prepared by PTs, the least preferred context was found to be education. The efforts of PTs to include information about the context to solve the questions prepared in the task design assignments were also remarkable. Giving a contextual title, context-based attribute names can be given as examples to this. However, although information about the source of data was given in most of the tasks, no information was given about the source of the data in four tasks. In general, it can be said that the PTs considered importance of creating connections with the context in the tasks they prepared. While this result concurs with the results of Casey et al. (2020; 2021), it differs from the results of Bakogianni (2015). While Casey et al. (2021) stated that one of the strengths of the tasks designed by the PTs is the constant connections made with the context, Bakogianni (2015) revealed that statistical context is not only a difficult learning goal to reach but also an important teaching challenge for mathematics teachers. Although calculating statistical measures and constructing representations are important skills, it is emphasized that students should be given the opportunity to conduct research within contexts in order to gain statistical skills (Bargagliotti et al., 2020; Casey et al, 2020; Franklin et al., 2007).

When the prepared task design assignments created by PTs were examined in terms of the SPS, it was observed that most of the task design assignments included the entire SPS, and only three tasks did not include the stage of interpreting the results. When the designed task design assignments were examined in terms of their SPS levels, it was revealed that the tasks were mostly at Level B, followed by Level A. The number of the tasks prepared at Level C was found to be the lowest. It can be said that the prepared task design assignments contained statistical questions and provided guiding information about data collection. Preparing survey questions and questionnaires on how to collect data can be given as an example to this. In different studies (Casey et al., 2020), attention was drawn to the paucity of tasks that prompted one to consider how data were collected or how the collection method might affect their interpretation. However, it is emphasized that teachers should support students to understand the data collection process (McClain & Cobb, 2001). In addition to enabling the exploration of different forms of representation, the CODAP program may have supported PTs in the process of analyzing the data by allowing them to perform calculations to visualize relationships between variables (e.g., the equation for the least-squares line). Many studies (e.g, Casey et al., 2020; Prodromou, 2015) have drawn attention to the convenience that different technological software provides in the statistics teaching.

In the process of interpreting the obtained data, interpretations were made on the basis of reading the data and reading between the data in Level A tasks, in addition to these interpretations, reading beyond the data was also performed in Level B tasks and in Level C tasks, reading behind the data was also performed. In general, reading the data, reading between the data and reading beyond the data interpretations were mainly performed in the task design assignments. In studies (Casey et al., 2020; Chick & Pierce, 2008), it is pointed out that the interpretations made by PTs are at a simpler level (reading the data, reading between the data) and that more advanced levels should be included. In this sense, recent results (Casey et al., 2021) and the findings of the current study can indicate that there are improvements in the quality of statistical tasks.

The fact that PTs mainly prepared task design assignments at Level A and B shows that they are limited in preparing task design assignments at Level C. There might be two reasons why PTs are limited in preparing tasks at Level C. Lack of knowledge of PTs may have caused them to have difficulties in preparing task design assignments at Level C. Studies (e.g. Burgess, 2007; Casey & Wasserman, 2015; Hannigan et al., 2013) draw attention to the fact that the lack of knowledge of teachers and PTs affects their task preparation skills. Another reason for the difficulties they experienced in preparing tasks at Level C may be the objectives in the curriculum in Turkey. It was explained to PTs that while preparing task design assignments, they could receive support from the curriculum as well as academic resources. Studies conducted (Batur, et al., 2021) have determined that the curriculum in Turkey mainly includes Level A objectives, followed by Level B objectives. There are no objectives at Level C. Furthermore, Balcı (2023), who examined textbooks and curriculum in Turkey, concluded that the step of interpreting findings in the curriculum is only limited to making interpretations and decisions based on evidence obtained from data analysis and does not explicitly include making inferences.

The results showed that although the PTs had some difficulties in the preparation of task design assignments (for example, asking mathematical questions, having few tasks at Level C), in general, their efforts to consider SPS task preparation components attracted attention. It can be thought that the PTs' discussions on how effective statistics teaching should be during the course they took, as well as their examining sample tasks, led to such a result. The fact that the CODAP software allows for the creation of multiple and various graphs and the calculation of various measurements (e.g., central tendency, dispersion) easily may have helped the PTs focus more on other components of the tasks (e.g., context, interpreting the results). Recent studies (Gorman, 2017) have shown that teachers tend to create their own teaching materials rather than using textbooks. When evaluated in this context, it is seen to be important for effective statistics teaching that PTs, who will be the teachers of the future, make progress in preparing SPS tasks.

The current study focused on the PTs' preparation of task design assignments. What was done in the tasks was based on guessing the thoughts of the students and the tasks were not implemented in the classroom environment. This can be considered as a limitation of the study. In future studies, it can be discussed how these tasks are reflected in the implementation process. Moreover, opportunities can be provided for PTs to work on large, multivariate and real data to design higher-order tasks and to experience a SPS by using software such as CODAP (Casey et al., 2020) so that they can prepare tasks at Level C.

References

Australian Curriculum, Assessment and Reporting Authority (ACARA) (2015) *Australian curriculum: Mathematics: Sequence of content*. https://australiancurriculum.edu.au/media/3680/mathematics_-_sequence_of_content.pdf

Bakogianni, D. (2015). Studying the process of transforming a statistical inquiry-based task in the context of a teacher study group. In K. Krainer & N. Vondrova (Eds.), *Proceedings of the 9th Conference of the European Society for Research in Mathematics Education (CERME9).* (pp. 615–621). Charles University.

Balcı, M. (2023). Teaching statistics in middle school mathematics: Investigation of the curriculum and course materials. [Unpublished master dissertation]. Hacettepe University.

Bargagliotti, A., Franklin, C., Arnold, P., Gould, R., Johnson, S., Perez, L., & Spangler, D. (2020). *Pre-K-12 Guidelines for Assessment and Instruction in Statistics Education (GAISE) report II*. VA: American Statistical Association.

Batur, A., Özmen, Z. M., Topan, B., Akoğlu, K., & Güven, B. (2021). A cross-national comparison of statistics curricula. *Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12*(1), 290–319. https://doi.org/10.16949/turkbilmat.793285

Ben-Zvi, D. (2011). Statistical reasoning learning environment. *Revista de Educação Matemática e Tecnológica Iberoamericana, 2,* 1–13. https://doi.org/10.36397/emteia.v2i2.2152.

Boaler, J., & Levitt, S. (2019, October 23). Modern high school math should be about data science-not Algebra 2. *Los Angeles Times*. https://www.latimes.com/opinion/story/2019-10-23/math-high-school-algebra-data-statistics

Braswell, J. S., Dion, G. S., Daane, M. C., & Jin, Y. (2005). *The Nation's Report Card [TM]: Mathematics, 2003*. NCES 2005-451. National Center for Education Statistics.

Burgess, T. (2007). Investigating the nature of teacher knowledge needed and used in teaching statistics. [Unpublished doctoral dissertation]. Massey University.

Burgess, T. A. (2011). Teacher knowledge of and for statistical investigations. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics – Challenges for teaching and teacher education* (pp. 259–270). Springer.

Carver, R., Everson, M., Gabrosek, J., Horton, N., Lock, R., Mocko, M., ... & Wood, B. (2016). *Guidelines for Assessment and Instruction in Statistics Education (GAISE) college report 2016*. American Statistical Association.

Casey, S. A., Harrison, T., & Hudson, R. (2021). Characteristics of statistical investigations tasks created by preservice teachers. *Investigations in Mathematics Learning, 13*(4), 303–322. https://doi.org/10.1080/19477503.2021.1990659

Casey, S., Hudson, R., Harrison, T., Barker, H. & Draper, J. (2020). Preservice Teachers' Design of Technology-Enhanced Statistical Tasks. *Contemporary Issues in Technology and Teacher Education, 20*(2), 269–292.

Casey, S. A., & Wasserman, N. H. (2015). Teachers' knowledge about informal line of best fit. *Statistics Education Research Journal, 14*(1), 8–35.

Chick, H., & Beswick, K. (2018). Teaching teachers to teach Boris: A framework for mathematics teacher educator pedagogical content knowledge. *Journal of Mathematics Teacher Education 21*(5), 475–499. https://doi.org/10.1007/s10857-016-9362-y

Chick, H. L. & Pierce, R. U. (2008). Teaching statistics at the primary school level: Beliefs, affordances, and pedagogical content knowledge. In C. Batanero, G. Burrill, C. Reading & A. Rossman (Eds.), *Joint ICMI/IASE study: Teaching statistics in School mathematics. Challenges for teaching and teacher education. Proceedings of the ICMI Study 18 and 2008 IASE round table conference*. ICMI/IASE.

Cobb, P & McClain, K. (2004). Principles of Instructional design for supporting the development of students' statistical reasoning, In D Ben-Zvi & J. Garfield (Ed.) *The challenge of developing statistical literacy, reasoning and thinking* (pp.375–396), Kluwer.

Cobb, G. W. & Moore, D. S. (1997). Mathematics, statistics, and teaching, *The American mathematical monthly, 104*(9), 801–823.

Concord Consortium (2019). *Common online data analysis platform (CODAP)* [computer software]. https://codap. concord.org/

Curcio, F. R. (1989). *Developing graph comprehension. Elementary and middle school activities.* National Council of Teachers of Mathematics.

Curcio, F. (1987). Comprehension of mathematical relationships expressed in graphs. *Journal for Research in Mathematics Education, 18*(5), 382–393. https://doi.org/10.2307/749086 da Ponte, J. P. (2011). Preparing teachers to meet the challenges of statistics education. In C. Batanero, G. Burill, C. Reading (Eds.), *Teaching statistics in school mathematics-Challenges for teaching and teacher education* (pp. 299–309). Springer.

delMas, R. (2004). A comparison of mathematical and statistical reasoning. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 79–95). Kluwer Academic Publishers.

Dierdorp, A., Bakker,A. Eijkelhof, H. & Maanen, J. (2011). Authentic practices as contexts for learning to draw inferences beyond correlated data, *Mathematical Thinking and Learning, 13*(1–2), 132-151. https://doi.org/10.1080/10986065.201 1 .53829

Franklin C., Bargagliotti A. E., Case C. A., Kader G. D., Schaeffer R. L., Spangler D. A. (2015). *The statistical education of teachers*. VA: American Statistical Association.

Franklin, C., Kader, G., Mewborn, D., et al. (2005) *Guidelines for Assessment and Instruction in Statistics Education. (GAISE) Report: A pre-k-12 Curriculum Framework*. American Statistical Association.

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Schaeffer, R. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) Report: A Pre-K-12 curriculum framework*. American Statistical Association, Alexandria.

Friel, S., Bright, G., and Curcio, F. (1997). Understanding Students' Understanding of Graphs. *Mathematics Teaching in the Middle School, 3*(3), 224–227.

Garfield, J. (1995). How students learn statistics. *International Statistical Review, 63*, 25–34. https://doi.org/10.2307/1403775

Gorman, N. (2017, February 7). Survey finds teachers spend 7 hours per week searching for instructional materials. *Education World*. Retrieved from www.educationworld.com.

Groth, R. E. (2007). Toward a conceptualization of statistical knowledge for teaching. *Journal for Research in Mathematics Education, 38*(5), 427–437. https://www.jstor.org/stable/30034960

Groth, R. E. (2013). Characterizing key developmental understandings and pedagogically powerful ideas within a statistical knowledge for teaching framework. *Mathematical Thinking and Learning, 15*, 121–145. https://doi.org/10.1080/10986065.2013.770718

Hannigan, A., Gill, O., & Leavy, A. M. (2013). An investigation of prospective secondary mathematics teachers' conceptual knowledge of and attitudes towards statistics. *Journal of Mathematics Teacher Education, 16*(6), 427–449. https://doi.org/10.1007/s10857-013-9246-3

Jones, D. L., & Jacobbe, T. (2014). An analysis of the statistical content of textbooks for prospective elementary teachers. *Journal of Statistics Education, 22*(3), 22–40 https://doi.org/10.1080/10691898.2014.11889713

Jones, D. L., Brown, M., Dunkle, A., Hixon, L., Yoder, N., & Silbernick, Z. (2015). The statistical content of elementary school mathematics textbooks. *Journal of Statistics Education, 23*(3), 1–22. https://doi.org/10.1080/10691898.2015. 11889748

Langrall, C. W., Makar, K., Nilsson, P., & Shaughnessy, J. M. (2017). Teaching and learning probability and statistics: An integrated perspective. In J. Cai (Ed.), *Compendium for research in mathematics education* (pp. 490–525). VA: National Council of Teachers of Mathematics.

Leavy, A. & Frischemeier, D. (2022). Developing the statistical problem posing and problem refining skills of prospective teachers. *Statistics Education Research Journal, 21*(1), 1–27. https://doi.org/10.52041/serj.v21i1.226

Lee, H. S. (2019, May 16). Data science education in 6-12 classrooms: What should a coulda woulda, but often ain't there. Presentation at the *United States Conference on Teaching Statistics*, State College, PA. Retrieved from: https://www.youtube.com/watch?v=53WuS5z30PY&feature=youtu.be&t=596

Lee, H. S., Kersaint, G., Harper, S. R., Driskell, S. O., Jones, D. L., Leatham, K. R., Angotti, R. L. & Adu-Gyamfi, K. (2014). Teachers' use of transnumeration in solving statistical tasks with dynamic statistical software, *Statistics Education Research Journal, 13*(1), 25–52.

Lovett, J.N. & Lee, H.S. (2018) Preservice Secondary Mathematics Teachers' Statistical Knowledge: A Snapshot of Strengths and Weaknesses, Journal of Statistics Education, 26(3), 214–222. https://doi.org/10.1080/10691898.2018.1496806

McClain, K., & Cobb, P. (2001). Supporting studentsí ability to reason about data. *Educational Studies in Mathematics, 45*(1/3), 103–129. https://doi.org/10.1023/A:10138745 14650

Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation*. Wiley Publications.

Ministry of National Education [MoNE], (2018). Middle school mathematics curriculum (Elementary and middle schools 1,2,3,4,5,6,7 and 8 grades). [Matematik dersi öğretim programı (ilkokul ve ortaokul 1,2,3,4,5,6,7 and 8. Sınıflar]. Ankara.

Neto, S. C. (2017). Combining distance and traditional learning: A study of the use of virtual learning environment objects and massive online open courses in statistics class. *International Journal of Information and Education Technology, 7*(1), 1–5. https://doi.org/10.18178/ijiet.2017.7.1.831

Peck, R., Gould, R., Miller, S., & Zbiek, R. (2013). *Developing essential understanding of statistics for teaching mathematics in grades 9-12*. VA: National Council of Teachers of Mathematics.

Perkowski, D. A., & Perkowski, M. (2007). *Data and probability connections: Mathematics for middle school teachers*. NJ: Pearson Prentice Hall.

Prodromou, T. (2015). Teaching statistics with technology, *Australian Mathematics Teacher, 71*(3), 32–40.

Putney, L. G. (2010). *Case study*. In N. Salkind (Ed.) Encyclopedia of research design. (pp. 89–103). Thousand Oaks, CA: Sage Publications.

Rossman, A., Chance, B., & Medina, E. (2006). Some important comparisons between statistics and mathematics, and why teachers should care. In G. F. Burrill, & P. C. Elliott (Eds.), *Thinking and reasoning about data and chance: Sixty eighth year book* (pp. 323–333). VA: NCTM.

Scheaffer, R. L. (2006). Statistics and mathematics: On making a happy marriage. In G. F. Burrill, & P. C. Elliott (Eds.), *Thinking and reasoning about data and chance: Sixty eighth year book* (pp. 309–322). VA: NCTM.

Shaughnessy, J. M. (2007). Research on statistical learning and reasoning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 957–1009). NC: Information Age Publishing.

Shaughnessy, J. M., Garfield, J. & Greer, B. (1996). Data handling. In A. J. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Ed.), *International handbook of mathematics education* (pp. 205–237). Kluwer Academic Publishers.

Shapiro, E. J., Sawyer, A. G., Dick, L. K., & Wismer, T. (2019). Just what online resources are elementary mathematics teachers using? *Contemporary Issues in Technology and Teacher Education, 19*(4), 670–686.

Suhermi, S. & Widjajanti, D. B. (2020). What are the roles of technology in improving student statistical literacy?. *Journal of Physics: Conference Series, 1581*(1), 012067. https://doi.org/10.1088/1742- 6596/1581/1/012067

Tishkovskaya, S. & Lancaster, G. A. (2012). Statistical education in the 21st century: A review of challenges, teaching innovations and strategies for reform. *Journal of Statistics Education, 20*(2), 1–56. https://doi.org/10691898.2012.11889641

Tran, D., & Lee, H. S. (2015). Considerations for design and implementation of statistics tasks. In *Teaching statistics through data investigations MOOC-Ed, Friday Institute for Educational*

Innovation: NC State University, Raleigh, NC. Retrieved from http://ficourses.s3.amazonaws.com/tsdi/unit_3/CDIST.pdf

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review, 67*(3), 223–265. https://doi.org/10.1111/j.1751-5823.1999.tb00442.x

Wild, C. J., Utts, J. M., & Horton, N. J. (2018). What is statistics? In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International Handbook of Research in Statistics Education* (pp. 5–36). Springer.

Weiland, T. (2019). The contextualized situations constructed for the use of students by school mathematics textbooks. *Statistics Education Research Journal, 18*(2), 18–38. https://doi.org/10.52041/serj.v18i2.138

Computer-based assessment in mathematics: Issues about validity

Anneli Dyrvold and Ida Bergvall

Department of Education, Uppsala University, Sweden

Computer-based assessments is becoming more and more common in mathematics education, and because the digital media entails other demands than paper-based tests, potential threats against validity must be considered. In this study we investigate how preparatory instructions and digital familiarity, may be of importance for test validity. 77 lower secondary students participated in the study and were divided into two groups that received different instructions about five different types of dynamic and/or interactive functions in digital mathematics items. One group received a verbal and visual instruction, whereas the other group also got the opportunity to try using the functions themselves. The students were monitored using eye-tracking equipment during their work with mathematics items with the five types of functions. The result revealed differences in how the students undertook the dynamic functions due to the students' preparatory instructions. One conclusion is that students need to be very familiar with dynamic and interactive functions in tests, if validity is to be ensured. The validity also depends on the type of dynamic function used.

ARTICLE DETAILS

LUMAT Special Issue
Vol 11 No 3 (2023), 49–76

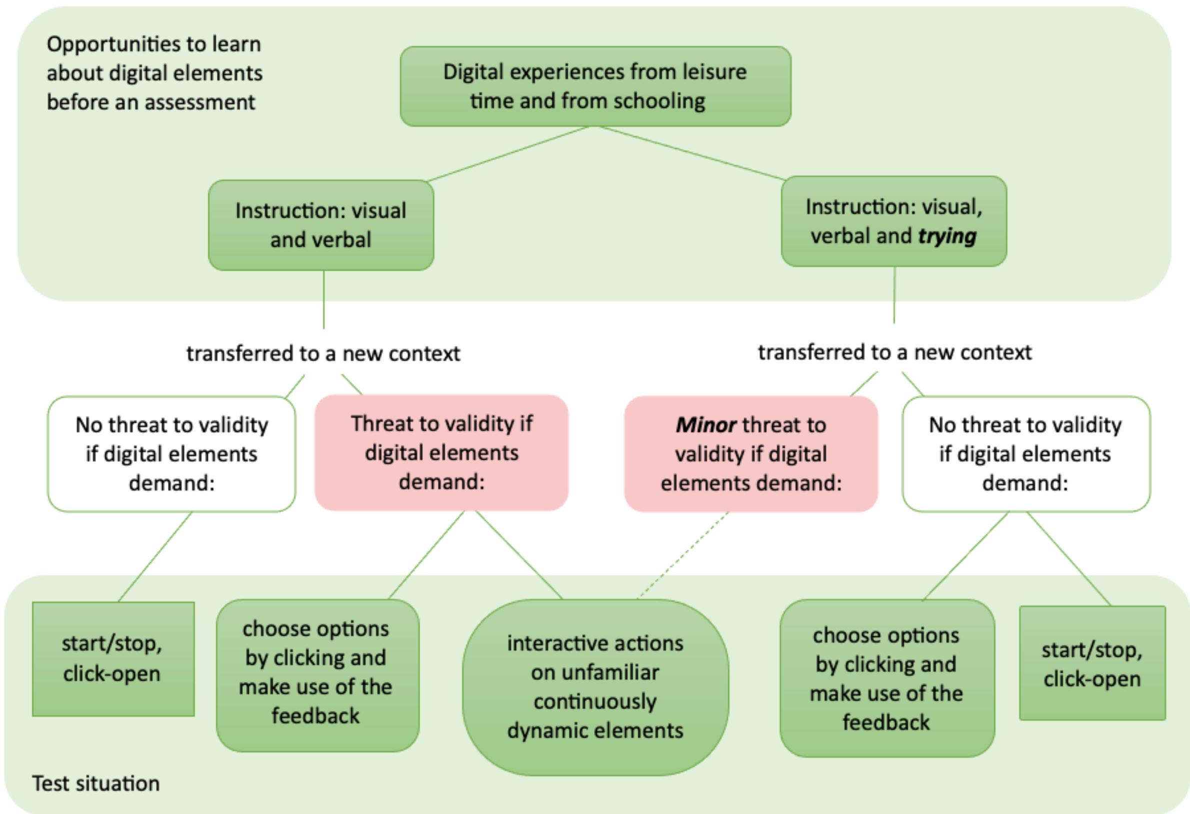
Received 27 October 2022
Accepted 14 September 2023
Published 5 October 2023

Pages: 28
References: 39

Correspondence:
a.dyrvold@gmail.com

https://doi.org/10.31129/LUMAT.11.3.1877

Keywords: computer-based assessment, dynamic, interactive, validity, transfer



 | CC BY 4.0

1 Introduction

The use of computer-based assessments is continuously increasing, in mathematics education as well as in other subjects. Computers can be used in tests to simplify data collection or because computers offer practical tools in the test situation (e.g., language editing) but the purpose can also be to test aspects of digital competence. The discrepancy between merely using computers as a data collection tool or to, more or less, also assess digital competence has implications for the kind of preparation or training that is needed before a test. In both cases it is crucial that the test-taker is not disadvantaged due to misunderstandings that prevent them from using a particular digital component in the test, something that can threaten the validity of a test.

There are many reasonable arguments for digitization of tests and an increase in digital tests are expected. Bennett (2015) points out that the evolution from paper-based assessments to electronic ones is substantive. Currently, different assessments are at different stages in this evolution. In the mathematics part of PISA (Programme for International Student Assessment), digital items are included based on the argument that "a level of competency in mathematical literacy in the 21st century includes usage of computers" (OECD, 2013, p. 44). It is emphasised in the PISA framework that the digital environment provides opportunities to include more interactive and authentic items, for example with drag-and-drop and with real-world data (OECD, 2013). The benefits of a dynamic environment have also made an imprint on the mathematics assessment in TIMSS (Trends in International Mathematics and Science Study). It is argued that computer-based assessments enable the incorporation of new and better assessment methods, through the use of digital components (Mullis & Martin, 2017).

The variety of skills demanded to navigate in a digital environment risk to be underestimated, in particular if familiarity with such an environment is taken as a guarantee for a broadly applicable digital competence. The younger generations have more or less lived their whole lives using digital devices as natural tools in their everyday lives, which is why these generations are called "digital natives" (e.g., Prensky, 2001). Digital natives are perceived as possessing advanced knowledge of the use of digital equipment and technology. This generalisation has however been questioned by many researchers (e.g., Bennett et al., 2008a; Helsper & Enyon, 2010) who highlight the great variation within the generations and that there are no direct differences between digital natives and older generations, even if digital natives use digital technology to a large extent. The use of digital devices in leisure time and professionally, or in schooling can be very different things and it cannot be taken for granted that the use of digital tools, for example in an assessment context, is problem-free.

In terms of test validity, it is crucial that students are comfortable using the digital functions offered in a test, which is not guaranteed even if they belong to a younger, "digital native" generation. It is of interest to explore what kind of familiarity with digital materials is needed for students to make us of use different kinds of digital functions in tests, because with such an understanding we can better prepare test takers, and thereby diminish threats to validity. Accordingly, this study addresses validity in relation to computer-based tests, with a particular focus on different digital functions and which kind of instructions that are sufficient before a test.

1.1 Aim and research question

The aim of this study is to contribute to understanding the role of students' acquaintance with digital functions for how they encounter digital teaching materials and tests. The research question is: In a digital environment, are there any differences between which dynamic functions students are prone to undertake and use in the face of different instructions proposed?

2 Background

2.1 Utilising affordances of the digital media in assessments and in teaching materials

Many tests in mathematics that previously have been offered in print are successively replaced by digital counterparts; for example recent versions of PISA mathematics as well as TIMSS mathematics are offered both in paper and in digital format (OECD, 2021; Mullis & Martin, 2017). There are a variety of reasons for using computer-based assessments. For example, digital resources have the potential to enrich items in tests by inclusion of new multimodal resources such as video explanations, hints or worked out answers, or mathematical software for manipulating objects like, graphing, drawing or solving equations (e.g., see Usiskin, 2018). Digital resources also provide opportunities to organise mathematical information in new ways, for example by linking to explanations and definitions that can be shown or hidden (O'Halloran, Beezer, & Farmer, 2018). Computer-based assessments may also facilitate formative assessment (e.g., Aldon & Panero, 2020; Barana et al., 2021) or contribute benefits related to the implementation itself, such as simplification of distribution, reduced working time for assessment and grading, automatic reporting of test results, increased usability and accessibility for students with disabilities through technical solutions (e.g., Regeringen, 2017; College Board, 2022). Results from comparisons between paper-based and computer-based mathematics tests have revealed that computer-based tests are significantly harder than the paper-based test (Bennett et al., 2008b) but there are also contradictory results. A study contrasting paper-based and computer-based tests in Korean language, mathematics, social sciences, and science revealed higher results on all paper-based tests, but only with a significant difference for Korean language. Furthermore, for female participants, the difference was significant also for mathematics (Hanho, 2014). The difference between modes (paper or computer) is also evident in students' choices of mode for responses. An analysis of which format students prefer to use when responding on tests reveal that the paper format is preferred by 71 percent (100 participants) when contrasted to digital pen or type-written response format (Davis et al., 2021). The mean results were also highest for the paper format and lowest for type-written response, but these differences were not significant. Accordingly, the occurrence of a mode effect seems to be dependent on the subject, but contextual factors such as familiarity with the hardware can also affect performance (Dadey et al., 2018).

There undoubtedly are numerous assets to computer-based tests, enabling a valid assessment of students' mathematics skills, but it is important to keep in mind that the opportunities provided by digital media are likely to place new demands on students' ability to read and navigate the digital environment and to work with the digital resources. It has been shown, for example, in a study of students' work with GeoGebra in ordinary teaching situations, that students with lower achievement levels in mathematics but used to working with GeoGebra outperformed high-performing students who were not used to working with GeoGebra (Baccaglini-Frank, 2021). This probably applies to assessment situations as well. Even when simpler tools such as a ruler or protractor are integrated digitally, using them demands other skills than if the physical tools are used. If such tools are integrated in tests, students must be accustomed to using them to perform well (Lemmo, 2021). Harris et al. (2021) showed that students' performance on interactive tasks was lower than on static tasks in the digital environment. The reduced performance may be due to increased cognitive demands on the students as they need to work with several different mental processes simultaneously in interactive tasks. The vertical order in which text elements are usually arranged on ordinary sheets also works less well on a screen with its proportions, and tasks may therefore need to be structured differently, for example in columns. This restructuring may complicate reading because the student needs to coordinate information from different parts of the text in a way that they may not be used to (Lemmo, 2021).

2.2 The relation between experience with digital teaching material and performance on computer-based tests

A relation between whether digital teaching material is used in training and performance on a test in digital form is much expected, but exactly which kinds of digital features test takers need experience of is not self-evident. In a context where test takers are accustomed to digital devices both in school and at home, a proficiency in navigating in the digital environment might be expected due to the familiarity with digital devices. There are previous studies pointing to the connection between experience of being taught with digital learning materials and success in computer-based tests. For example, Hamhuis and colleagues (2020) explored potential differences in performance of Dutch primary school students on the TIMSS test, depending on whether it was paper based or conducted on a tablet. The result revealed no significant differences between the paper and the tablet tests, which can be explained by the fact that Dutch students are used to the digital format. The picture is somewhat complicated by other studies such as Smolinsky et al. (2020) showing that university students who were taught with paper-and-pencil have slightly better results on computer-based tests compared to students who were taught entirely computer-based. This result speaks against the fact that general digital habits make it easier for students when working with computer-based tests. Within another research project (Hoch et al., 2018), an interactive mathematics teaching material for the introduction of fractions was designed and evaluated. The material consisted not only of digitised text, but also of three key interactive features; interactive exercises, adaptive demands and automatic feedback. A result that emerged from the study was a negative relationship between the time students spend on the tasks and performance (Hoch et al., 2018). Hoch and colleagues suggest that an explanation to the result is that the students did not use the exercises only to acquire new knowledge, but did also spend time practising rational number concepts. Time was therefore not a good measure of proficiency, even though practising is of course a desirable activity.

Students who are used to working with digital teaching materials and various types of dynamic resources in their daily lessons are likely to have advantages when encountering new types of resources in test situations in contrast to those using print material. Research has however shown that in contemporary teaching material comprising whole courses, the utilisation of highly dynamic elements is still rather limited (e.g., see Dyrvold, 2022; Ilovan et al., 2018). Students who study with digital textbooks can therefore also meet unexpected demands in computer-based assessments.

2.3 Validity in computer-based assessments

A crucial question in relation to computer-based assessment is of course what the test aims at assessing. It can be argued that digital competence is part of a comprehensive mathematical competence (e.g., see Geraniou & Jankvist, 2019) and tests can accordingly aim at assessing also mathematical digital competence. However, such an aim is not the goal in all digital mathematical assessments, but even if that was the case, it needs to be scrutinised which kinds of digital demands are wanted in tests and not. This issue relates to Messick's (1995) argumentation concerning the essentiality of construct validity, and in the design of computer-based tests, it is decisive whether the ability to handle digital tools is part of the construct or not. Messick highlights the complexity of capturing the construct a test aims at assessing. A test measure is only one indicator of the full construct, and it is fundamental that this indicator constitutes a good representation of this construct; that is, whether task responses are solely dependent on the processes, knowledge, and strategies utilised in the performance. If the intention is not to assess digital competence, the test should not reward digital abilities, nor should the test result be affected by unwanted digital difficulties.

Another potential threat to validity is construct irrelevant variance, which can be due to construct irrelevant easiness or construct irrelevant difficulty. An example of irrelevant easiness is when the possible answers in a multiple-choice question provide 25 percent chance of answering correctly by guessing, and irrelevant difficulty might be if dynamic resources in a task are difficult to undertake, so that students fail on tasks that they would otherwise be able to solve. Computer-based assessments increase the potential sources of construct irrelevant variance because the digital environment may entail skills not demanded in paper-based tests. Crucial for validity is therefore whether digital skills needed in the digital environment, is regarded as a part of the construct, or not. In this paper, we intend to contribute to understanding validity in computer-based tests by investigating the role of students' acquaintance with digital functions for how they encounter digital teaching materials and tests.

Regarding the mathematics content included in the construct of a test, Ripley (2009) distinguishes between computer-based tests designed based on a *migratory* strategy (keeping the test similar to the print version) or on a *transformative* strategy (inferring demands of new digital skills in the test). Depending on strategy, different validity issues arise. In assessments designed with a migratory strategy, effort is laid on ensuring the assessment is kept as similar to the print version as possible and several studies have proven a high validity in this type of assessments (Junpeng et al. 2019; Hamhuis, 2020). These studies indicate that digitalisation can be made without decreasing validity, but if many new functions are used in a computer-based test the risk of validity issues increases. Assessments based on a transformative strategy, for example, that are designed to bring *innovation* in curriculum design and learning introduces many elements that may be new to the test takers. An innovative assessment can provide rich opportunities to assess comprehensive mathematical competences that would be harder to capture in a paper-based test. The test taker can for example be offered dynamic materials (Yerushalmy & Olsher, 2020) and the tasks can contain real-world datal and present students with an authentic, simulated environment (e.g., see OECD, 2013). Bennett (2015) distinguishes assessments with a transformative strategy in a *second* and *third* generation. The second generation of assessments use new item formats including for example multimedia or constructed response options and may aim at assessing new constructs. The third generation is defined as assessments that uses complex simulations and interactive features and serves both individual and institutional purposes. These kinds of assessments are integrated with instruction with repeated sampling, and accordingly new skills are assessed in more sophisticated ways. Bennett concludes that challenges in relation to the third most advanced step in this evolution of digital tests is large and that the need for a cautious development process is essential. The current study addresses features that are prominent in the second and third generation of assessments; the inclusion of multimedia as well as interactive features. The many benefits of the third generation of assessments are addressed by Bennett but some important challenges are also highlighted. In particular he points out that the most important challenge to address in research is validity and fairness of the third generation of assessments for all individuals, in especially for students at risk.

The need for awareness of potential threats to validity when using innovative assessments is essential. For example, if the functions used are novel to the test taker there is a risk for construct irrelevant difficulty and the integration of several types of skills in the construct being assessed can probably make it difficult to distinguish construct irrelevant variance. Messick (1995) exemplifies how demand of communication skills can be judged either as irrelevant in an assessment of mathematical knowledge or considered as parts of mathematical proficiency and therefore relevant to the construct. For assessments that are complex in the sense of subsuming multiple processes, awareness of construct irrelevant variance is of particular importance. The sources of variance need to be thoroughly evaluated in relation to the construct being assessed, judging how compelling the evidence for the relevance of some variance is. It is apparent that the risk for construct irrelevant variance increases as new affordances of the digital media are used but the threats to validity can also decrease if for example the option to write with digital tools increases a test taker's ability to represent their knowledge. In addition, there are many affordances of the digital media to take advantage of besides distribution, central grading and curriculum innovation. Other possibilities could be to include certain dynamic functions, such as possibilities to write more detailed answers, to use digital functions or to include, for example, GeoGebra, but without aiming at curriculum innovation. Especially if computer-based tests would be designed in accordance with digital learning materials, the potential for meaningful and valid assessments is substantial.

The current study addresses validity issues by putting the focus on how test takers are prepared for a computer-based assessment; whether barely instruction is sufficient or if the test takers might need to try features that they are expected to utilise in the test.

3 Method

The current study is part of a larger project about digital teaching material in mathematics. In this study we focus on students' encounters with different dynamic elements in mathematics items after receiving one of two kinds of instruction about the different elements. At the core of the study is differences between types of dynamic elements and between two kinds of instruction given to the students before problem solving. Thorough descriptions about the elements and the instruction are therefore given in this section. Elements are defined as a coherent part of a text that may include both words, symbols, and images, where the constituents can be static or dynamic.

Some of the elements are also interactive. The types of dynamic elements focused in this study (Table 3) is utilised in the *Facts* in an item, see Figure 2.

3.2 Data

Empirical data in this study was collected in an eye-tracking analysis of 77 grade nine students from two different schools in Sweden, during work with mathematics items. Both schools used printed mathematics textbooks for regular teaching. Each participant took a survey with questions about digital habits before working with the items. Information about the participants' grades were also collected. Data from three participants were excluded due to bad calibration or missing data from the eye-tracking. One participant rushed through all items in four minutes and data from this participant was also excluded. Data from three students were excluded because these students had not reached the first level (grade E) in Swedish as a second language. The participants were divided in two groups, referred to as *Show* and *Try* because in addition to verbal instructions one group was shown information before doing the items whereas the other group also tried dynamic features such as to start a film. Six students studied Swedish as a second language and had reached at least grade E (pass). Data from these students were also analysed because they were equally distributed between the *Show* (7%) and *Try* group (8%). All students who had reached the age of 15, gave a written consent to participate in the study (or their guardians otherwise). The students were informed that the overall purpose of the study concerned work with digital teaching materials in mathematics, and that the analysis would be carried out using eye-tracking analysis to monitor the participants' work on five mathematical items. The students were also informed that all participants are de-identified and that they could withdraw their consent at any time.

Ideally the selection of participants in the *Show* group and the *Try* group would have been matched pairs, which was not the case because during data collection the two types of information presentation were used on different occasions in two different schools. Because the study was designed ad hoc within a larger project data for the *Show* group were collected first and for fewer participants. Information about the participants experiences of digital teaching material in mathematics were gathered though a survey that teachers at the schools completed (Table 1). The survey is designed to capture a generalised view of classrooms in the two schools, and the frequency the resources were used could be ticked at three levels. A guide defined "never" as *never or at one occasion*, "seldom" as a *few times per semester*, and "often" as *several times per semester, maybe every week*. The results reveal differences between the groups for questions 4, 5, and 7.

Table 1. Use of Digital Teaching Resources in Mathematics in the Show Group and the Try Group.

	Show group (*n*=24)		**Try group** (*n*=46)
If teachers use particular digital resources in classroom:			
1. Smart board or similar	seldom		seldom
2. Digital quizz	often		often
3. Software to dynamically present mathematics (e.g. GeoGebra)	seldom		seldom
If students use particular digital resources in math:			
4. Computer or padlet	seldom	≠	often
5. Watch short film individually	often	≠	seldom
6. Software to dynamically present mathematics (e.g. GeoGebra)	seldom		seldom
7. Mathematics apps (e.g. for repetition)	seldom	≠	often

Note. The teachers chose between alternatives: never, seldom, often.

Comments given in the survey explains the differences. Regarding question 4: In both show and the try group all student have a personal computer or padlet but computers are not used often in mathematics. The try group explains that ”often” refer to the use of computers to look up solutions online or to use a particular app (Magma, see https://www.magma.se/). Regarding question 7: The choice ”often” for the try group is explained by the use of one particular app (Magma) on a regular basis: ”at least all students have access to the app”. In summary the differences between the groups are explained by one group’s more frequent use of short films and the other group’s use of one particular app; therefore it can be concluded that the two groups’ experiences of digital teaching materials are fairly similar. The results from the survey only give a general view of the participants experiences of digital teaching material, but at least the results assures that the two groups are not offered very different experiences of digital teaching materials at their schools.

Background information and information from a survey the participants completed was also used to ensure that the two groups were not too different regarding qualities that may play a role for their ability to learn from the instruction and to do a computer-based test. The experimenters followed strict protocols during the data collections and the experimenters had the same role throughout all data collection to avoid differences else than those intended (Show vs Try). The background information about the participants is presented in Table 2. The information about activities students do more than 7 hour a week is based on options in a Likert scale: “not at all”, “1-3h”, 3-6h”, ”7-14h”, “more than 14h”. The participants answered the questions “On your leisure time, approximately how many hours per week do you spend” … “on computer games?” and “on other digital activities?”. An experimenter was available to answer questions. The two options with at least 7 hours a week are included in the share presented in Table 2.

One difference between the groups is that there is a larger share of girls in the *Try* group, but we have no reason to believe gender plays a large role in interpretation of information before solving the tasks. A comparison between the *Show* and the *Try* group reveals that a larger share in the *Show* group states that they spend more time on computer games and a larger share in the *Try* group does other digital activities. For the digital activities together, however, the share of students is for the two groups .74 and .77 respectively. These numbers of added fractions must be interpreted with caution because the same student can be represented in both the share who plays computer games and the share who does other digital activities.

Table 2. Information About Participants in the Two Groups of Students.

	Show group (*n*=24)	**Try group** (*n*=46)
Fraction of girls in group	.36	.52
Fraction who plays computer game ≥7h/week	.39	.30
Fraction who does other digital activities than gaming ≥7h/week	.36	.47
Mean grade in mathematics (lowest 0, highest 5)	2.43	2.71
Mean grade in Swedish (lowest 0, highest 5)	2.64	2.43

In Sweden a scale F–A is used in grading. F represents to not pass and E-A increasingly higher grades. A comparison of grades in the subjects Mathematics and Swedish revealed fairly similar mean grades between the groups. The *Show* group have slightly higher grades in Swedish and the *Try* group have slightly higher grades in mathematics.

3.3 Mathematics items with different dynamic features

Five different mathematics items were used. All students were presented these items in the same order: Item 1, 2, 3, 4, and 5. All items have some essential Facts that is needed to solve the task, at the right-hand side of the item (Figure 2). The Facts are designed in five versions for every item, based on a typology of elements designed to mirror an increasing interactivity and dynamics from type I to type V (Table 3). Accordingly there are five versions for each item. A counterbalanced combination of items in different versions was used, which ensured each student was offered all five items, and all five types of elements used in presentation of the Facts but in counterbalanced order. Thus, the same timelines with the counterbalanced order of items were used in the *Show* group and the *Try* group.

Table 3. Typology of Elements Used in Facts.

Element type	Dynamic and interactive characteristics
I	the constituents are presented similar to a printed counterpart, but on screen
II	the constituents appear after a click on a button
III	the constituents are presented in a film with a voice
IV	has constituents where the reader needs to choose options by clicking to receive a response and if needed try again and finally make use of the feedback
V	has constituents that change continuously over time when the reader drags or grab and move objects with the mouse

For all element types except the static (type I) and the film (type III) there are labels or instructions that inform the test taker about what is expected. In the static version no actions are needed by the test taker and the film is presented in a very familiar format with a triangle ▶ as the start button. In type II the button has the inscription “Click to open” and in type IV there were options to click on, a button “Check” and another button “Try again”. In type V there is a sign “Drag” with an arrow pointing to a coloured dot that should be dragged. Examples of the design of dynamic facts of element type IV and V are given in Figure 1.

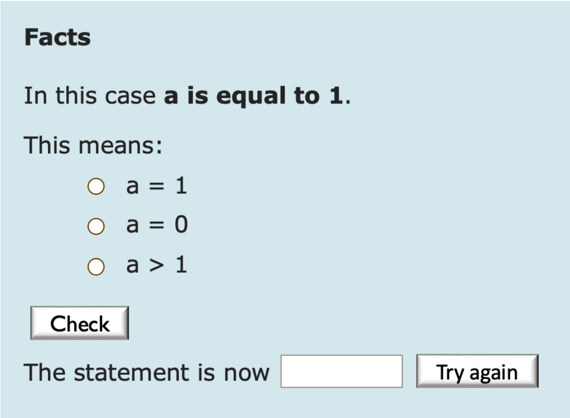
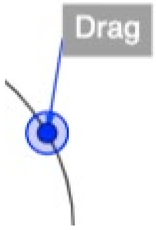
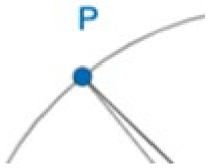
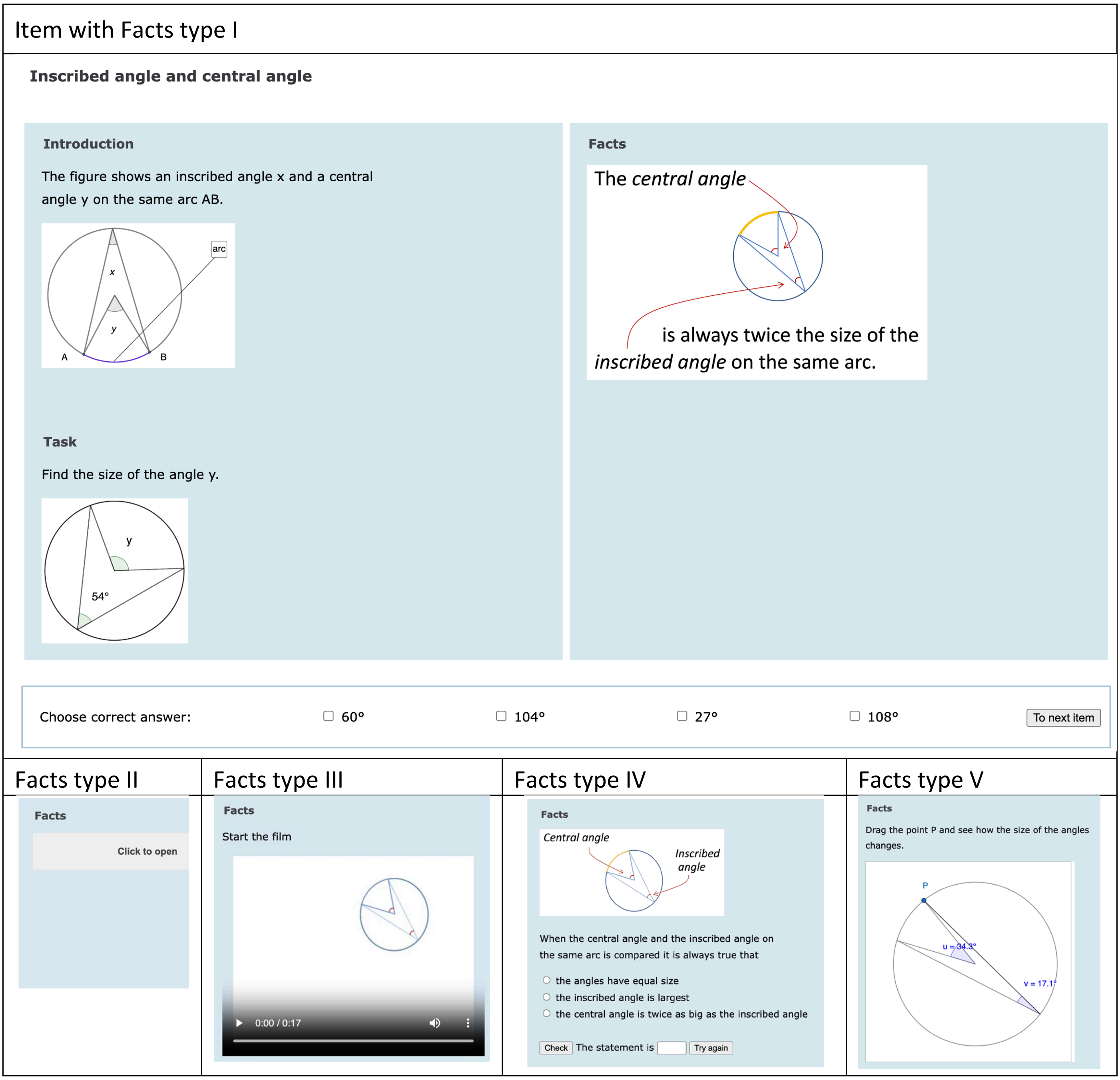
<table>
<tr><td>Type IV</td><td colspan="2">Type V</td></tr>
<tr><td>
</td><td>
</td><td>Drag the point P

</td></tr>
<tr><td>When an option is chosen and “Check” is clicked on, “True” or “False” is displayed. After clicking “Try again”, the user can start over, trying to find the correct option.</td><td colspan="2">The blue dot can be grabbed and dragged with the mouse.</td></tr>
</table>

Figure 1. Examples of Facts Expressed by Dynamic Element Type IV and Type V.

The items were chosen to be new to the participants so that the tasks could not be solved based on previous knowledge only; in particular the intention was that to solve the tasks the participants would need the information provided in the Facts. The mathematical content in five items is: 1) the inscribed angle theorem, 2) maximum and minimum of quadratic functions, 3) set theory, 4) the relation between power and roots, and 5) permutation and factorials. As an example, the item about the inscribed angle theorem and all five versions of Facts are presented in Figure 1. The exact same image and wording is used in Facts type I–III but in Facts II the Facts are displayed after a click on the button and in the film the text and image with arrows appear successively accompanied by a voice reading the text.



Note. The font of the different Facts is the same in all versions, here displayed in different sizes. In Facts V, the value for u disappears for u > 90

Figure 2. Example Item with Five Different Versions of Facts.

The development of the typology of elements and further examples of elements Type I–V can be found in Dyrvold (2022). This typology as well as the items were developed within a larger project and accordingly, the dynamic and interactive element types were not chosen to be demanding and test whether the participants chose to use them. In fact, the current study was developed ad hoc when participants choice to not use some Facts were identified. This omission was not expected and accordingly the current study was deemed important in an era with expansive use of digital assessments.

3.4 Implementation: Information to test takers

To diminish distractions caused by the eye-tracking equipment and to diminish extraneous cognitive load, all participants were individually informed about the setting and about the items before the test. This information was given just before they entered the room where they worked with the items. The information about the eye-tracking was very brief and the purpose was to answer questions and to cope with potential worries. The information about the items were strictly about the visual appearance of the items and about digital functions of the items. Both student groups received the same information about the eye-tracking and about the visual appearance of the items.

All students were shown a test item (Figure 3) on a screen and were told that the particular item also would appear as the first item in the test setting, before the five items included in the study. They were also told that the test item was intended to be easy and that it only played the role of an example. While looking at the item the students were informed that all items that should be solved had the same main arrangement. Thereafter the students were informed about the main parts in the items. This information was accompanied with gestures pointing at the parts on the screen. Information was given that all items have:

- a title,
- an introduction,
- some facts that are essential to solve the task,
- a task,
- answers to choose from, and
- a button that takes the student to the next item.

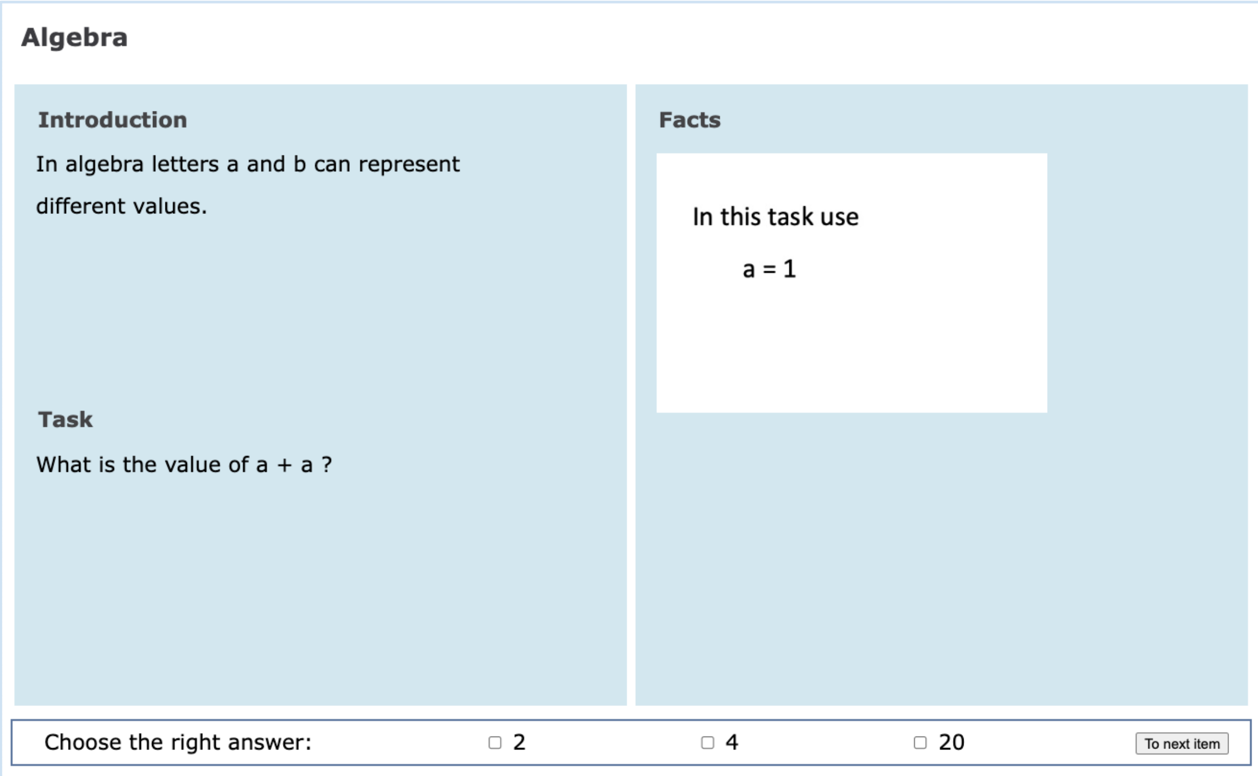


Figure 3. Test Item About Algebra with Facts in Static Form.

The use of a practice item as a first item in the test setting means data from the participants first minute, when getting comfortable with being tracked, do not affect data intended for analyses. Because all items have the same parts (Figure 2) present in the same spot on the screen, the reader does not need to grapple understanding the information flow, which is likely to reduce some extraneous cognitive load. The training with a very easy test item also contributes to making the setting more similar to the use of a digital teaching material that is familiar to the reader, as in the school setting.

After receiving the first information about the main parts in the items, the students were informed that different element types were used to present the Facts in the five items and that these Facts offered information essential to solve the task. Five versions of the test item, about algebra, were used to illustrate all different element types. This information was presented differently to the two groups. The *Show* group received the visual information from a PowerPoint presentation on a laptop. The *Try* group received the visual information from html files on a laptop, and for the four items with Facts that had some dynamic element the *Try* group were asked to try the dynamic element by clicking/dragging or likewise. The spoken information was given based on a list of information and was therefore the same for the *Show* and the *Try* group, with exception of information related to the different settings.

- For the first element type in static form, the information was the same for both groups. They were informed that these Facts were in static form, visible directly.
- For the second element type both groups were informed that one item had a button that must be clicked on to "open" the Facts. The *Try* group clicked on a button and the Facts appeared whereas the *Show* group was shown where the button would appear. For both groups, the button had the inscription "Click to open" as in the test setting.
- The third element type was in the form of a video. Both groups were informed that they could start the video and look many times if needed. They were also informed that headphones were used so no one (i.e. the experimenter) would reflect over whether they looked many times. The *Try* group was also encouraged to try and start a film.
- For the fourth element type both groups were informed that this type of presentation of the Facts was a bit like a task because they should choose an option and thereafter click "Check" to see if the option was correct. They were also informed that they should try again if they choose the wrong option. Finally, they were instructed not to think of this activity as *the task* but as contributing to make the Facts complete. They should still do the task. The *Try* group clicked on one option and then on "Check". The *Show* group were only shown a static page displaying the layout of the options and were informed that they were supposed to try different options, check whether it was correct and retry if needed (Figure 1).
- For the fifth element type both groups were informed that there would be some content that can be dragged or moved (Figure 1). The *Try* group moved a slider and saw a number appearing whereas the *Show* group were informed that there would be a prompt to drag and were shown how it could look visually. The experimenter showed a grab-drag gesture in front of the screen visualising how to drag a slider.

The participants were asked if they had any queries and some students asked questions, for example whether they should use a mouse or a trackpad.

3.5 Eye tracking

The participants' activities while working with the practice item and the five mathematics items with Facts expressed using five different element types were captured with an eye-tracking camera. The results from the eye-tracking data are mainly used in other studies, only a minor part of the data is used in the current study. Based on participants' mouse-clicks and eye-fixations on the Facts it was coded whether a participant used the Facts or not. The static Facts was coded as *used* if the participants had fixed their gaze on the constituents in the Facts at least once. For the Facts that have a dynamic component all items were first coded (tentatively) as used if the student had clicked on any button in the facts. To ensure that the participants for which the Facts were coded as *used* actually had used the Facts, screen recordings where students' gaze were displayed were also examined for all participants. Based on the recordings, all cases where participants that had clicked on a button or dragged a dot in the Facts were coded as "used". All participants that used the Facts like this, also spent time reading the Facts, which could be seen by the gaze-path on the Facts. The difference between the measure *gaze fixation* (for Fact I) and *used* (for Facts II-V) means that the static Facts have the role of a reference.

3.6 Transfer as a theoretical frame of analysis

Understanding information and being able to use the information in new situations means that the information needs to be correctly interpreted but also later retrieved and related to a new context (i.e. transfer). In the current study oral explanations are given to all participants whereas one of the two groups also had the opportunity to try and use the dynamic features themselves. Potential differences in outcome between the two groups are therefore likely to be related to the opportunity to learn by trying the different dynamic elements. The hypothesis is that trying supports the ability to transfer information from one situation to another. Theory about transfer provides a frame for the analysis in this study.

Transfer has traditionally been defined as the application of results from prior learning in novel situations (e.g., Gass & Selinker, 1983). At the core of the historical transfer perspective is Thorndike's identical elements (Thorndike & Woodworth, 1901). Identical elements refer to overlapping features between the learning situation and the new transfer situation. This narrow focus on identical elements has been criticised. One reason is that this focus implies that abstracted rule-like processes define successful transfer of the overlapping features and perceptual richness, for example when different forms of representation are used, is assumed to hinder transfer (e.g., Kaminiski, Sloutsky, & Hecksler, 2013). The concept of transfer has however been developed. Lobato & Hohense (2021) describes that the original, somewhat narrow, cognitivist perspective has been complemented with several other perspectives. For example, the actor-oriented transfer (AOT) perspective has been developed in response to limitations of previous conception of transfer. AOT includes the students' experience prior to the learning situation (e.g., see Lobato, 2012). Taking the actors view on transfer puts the lens on instances where prior experiences shape the students' activities in a transfer situation (Lobato, 2012). This means that this perspective is particularly useful for qualitative studies of how learners interpret situations and make connections (Lobato & Siebert, 2002). Despite the quantitative take in the current study, the AOT perspective is used to elucidate and discuss students' activities in the transfer situations and relate them to the students' previous experiences regarding gaming and other digital activities.

Furthermore, the AOT perspective puts focus on how *contextual sensitivity* can play a role for transfer (Lobato, 2012). Contextual sensitivity is defined as students' ability to utilise knowledge from previous experiences and based on the context adapt the knowledge to a variety of new transfer situations. In the current study, the transfer situation entails students' work on digital items on their own, items designed to minimise superfluous contextual factors. Thus, contextual sensitivity entails adaptation to a stringent context where the dynamic functions are the same as in the learning situation and a test situation where the students work on the digital items independently on their own.

In accordance with Nathan and Alibali (2021) and Goldstone et al. (2008) the role of perception and interactive processes are included as experiences that play roles in transfer. These processes are valuable to include because of the reciprocity between cognition and action (Nathan and Alibali, 2021). In the current study, students in two groups are offered different opportunities for perception and interactive processes when receiving instructions in the learning situation. The students receive instructions with or without the opportunity to try the dynamic functions and thereby experience rich perception and interactivity. In this way, the learning situations offer different modes to the two student groups. Due to the differences offered in the instructions in the learning situation, the ability to form and maintain connections between the learning and transfer situation, are supposed to be dependent on different demands on sensitivity to context.

For the analysis in this study, we use the notion of transfer to understand the process when students retrieve and transfer information from one context (the examples that the students were shown or offered to try before the test situation) to another context (the test situation). In this study, information refers to the function of the dynamic elements, not to the mathematical content in the items. Thus, successful transfer is expected to show by participants use of a particular element. It is assumed that the *Try* group is advantageous because the interactive preparation offered to this group facilitates the transfer and the students' contextual sensitivity, as well as it provides advantages depending on the relation between action and cognition.

3.7 Statistical analyses

Chi square tests were used to test whether the proportion of students who use the element types differ significantly depending on the kind of instruction given beforehand. Fisher's exact test was used in analyses where at least one cell in the chi square test had an expected count less than 5 and Pearson chi square with continuity correction was used if not. The reported p-values are for two-sided tests and $p<.05$ is used as significance level. The phi coefficient is calculated to analyse effect size and Cohen's criteria for effect size is used as a rule of thumb (Cohen, 1988).

4 Results and analysis

The analyses in this study were conducted to contribute to understanding the role of students' acquaintance with digital functions for how they encounter digital teaching materials and tests. The results reveal differences in the participants' actions depending on which instruction they had received. In particular, students tend to more frequently miss elements that are both dynamic and interactive, if they are not given the opportunity to try similar elements as part of the instruction. Analyses of the participants' interactions with items including the different dynamic functions are displayed in Table 4.

4.1 Results

In summary, because all participants used Facts presented using element type I (static), they appear to have understood that the role of the Facts was to offer information essential to solve the task. Very few participants missed the opportunity to use Facts presented with elements of type II-III. For Facts presented with element types IV and V on the contrary, there were many participants who did not use the Facts. There were significantly more participants in the group who only received instruction (*Show*) that missed the opportunity to use Facts presented using element type IV and the effect size was large. For Facts presented using element type V there was no significant difference between the groups; several participants in both groups did not use the Facts. Recall that element type I is in static form, element type II has a click to open function, element type III is presented as a film, element type IV is in the form of choices to click on options and retry until correct facts are presented, and element type V has a drag option that reveal continuous changes in the presented information (Figure 1).

Table 4. Whether Participants Use (No/Yes) Facts Presented With Different Element Types.

	Show group N=24		**Try group[1] N=46**		**Fisher's exact test**	
	no	yes	no	yes	p-value	Phi-coeff.
Facts I	0 (0%)	24 (100%)	0 (0%)	46 (100%)		
Facts II	2 (8.3%)	22 (91.7%)	3 (6.5%)	43 (93.5%)	1.00	.033
Facts III	2 (8.3%)	22 (91.7%)	0 (0%)	46 (100%)	.114	.237
Facts IV	9 (37.5%)	15 (62.5%)	0 (0%)	46 (100%)	**<.001**	.532
					Pearson chi square[2]	
Facts V	10 (41.7%)	14 (58.3%)	7 (17.9%)	32 (82.1%)	.077	. 259

[1] Data from seven participants were excluded from the analysis of Facts V because they could not see these facts or because no data were collected due to technical issues.
[2] Pearson chi square with continuity correction is used in this test because no cells in the chi-square test have an expected count less than 5.

Based on these results it can be concluded that students are not equally equipped for a computer-based assessment if they are only presented with information in contrast to if they also are given the opportunity to try the dynamic functions that are utilised in the assessment and use them themselves. It can also be concluded that the more interactive and dynamic functions that are included in an assessment the larger is the risk of missed opportunities to fully understand the offered information.

4.2 Analysis

An analysis of the results in relation to actor-oriented transfer elucidates the role of both previous experiences and of modes in the instruction, for the ability to utilise affordances of the dynamic elements in an item. The results differ between items with elements type I-III and type IV-V. No substantial differences between the groups (*Show* and *Try*) previous digital experiences or grades have been identified and it can therefore be assumed any differences in use of the elements are related to the two types of instruction. The element type I (static) plays the role of a standard type to contrast the results against and 100 per cent of the participants used that element. Elements type II-III (click to open, and film) are very similar to elements it can be assumed all participants have experiences of using. These elements are used by almost every participant in the study (95%). This result indicates that previous experiences, presumably from both leisure time and teaching, has provided the participants with sensitivity to context in use of these elements. Because there are no differences between the groups tendency to use the dynamic elements, it is likely the instruction in the learning situation plays a minor role for the use of the element type in the transfer situation, or at least that rich perception and interaction in the instruction is not needed. From a validity perspective this implies that in computer-based tests there is no substantial difference in the students' achievement due to instructions or previous digital experiences when element type II and III are being utilised in the test. This means that in environments similar to the Nordic school context these kinds of digital elements do not threaten the accuracy by which the test distinguishes between test takers based on their mathematical proficiency.

The largest threat against validity is however identified in relation to dynamic elements type IV and type V. For element type IV there is a significant difference in the use of facts ($p<.001$, Phi coeff.=.53) between the groups and for element type V the difference is nearly significant ($p=.077$). The conclusion is made because the results indicate that a thorough instruction that includes an opportunity to try the element is essential to assure transfer related to the use of those elements. The need for instruction is apparent for these elements and even with instruction (*Show* or *Try*), as many as 20 percent of the participants miss the opportunity to use at least one of the elements (IV & V). It can be assumed that the participants intend to make connections between previous experience and the transfer situation, but because the demands of interaction with the dynamic environment is large, knowledge from previous experiences may not be sufficient for successful transfer. Experiences from previous situations including dynamic and interactive elements likely differ both in features of the elements, and in the context they have been met, which can make transfer harder. Accordingly, when utilising dynamic elements that differ from those used in previous situations, the design of the instruction is eminent, and our results reveal that this is particularly important when the elements are highly dynamic and interactive as elements type IV and V.

The only element type that was used to a significantly different amount between the two instruction groups was type IV, the element where students are supposed to click, check, and potentially retry, to compose a correct mathematical statement and use that to solve the task. All participants in the *Try* group used the element in the transfer situation which means that the perceptually rich and interactive instruction were sufficient for successful transfer. On the contrary, 37.5 percent of the participants who did not get the opportunity to try elements in the instruction abandoned the option to use the dynamic element. This may be caused by differences between modes offered in the learning situation and the transfer situation, which can put too large demands on students' sensitivity to context in the target situation (the test). This result highlights a threat to validity and a reasonable source to it, namely instruction that does not provide a learning situation sufficient for the test takers to form and maintain connections between the learning situation and the transfer situation. An unwanted consequence of such instruction is construct irrelevant variance, leading to wrong inferences from an assessment.

5 Discussion

Computer based tests in mathematics are becoming more and more common as schools and teaching in general becomes more and more digitalised. Computer based tests make it possible to take advantage of features unique to the digital media, for example the inclusion of real-life data and various dynamic features (e.g., see Yerushalmy & Olsher, 2020; Ripley, 2009). Thus, innovative and computer-based tests can be used to assess other skills than those possible to test in paper-based tests; for example modelling competence may be easier to test accurately in a digital than a paper-based test. One purpose of the PISA-test for example, is to assess digital literacy, but most often the assessment is carried out with other intentions; an example is Swedish national tests where one aim is to ensure equal grading between different teachers, schools and principals. For high-stake assessments validity is, of course, of utmost importance, and an important issue with relevance for a sustainable mathematics education is how validity can be ensured.

We can assume that it is crucial that the students master the digital environment to ensure validity in a computer-based test and there is also evidence that experience with the digital medium can lead to better test results (e.g., Baccaglini-Franck, 2021; Dadey et al., 2018). There is also evidence that students must be accustomed to using dynamic functions to perform well (Lemmo, 2021; Harris et al., 2021). The current study contributes by highlighting the substantial risk to overlook or underestimate the need for apt instructions as preparation for a computer-based assessment. The use of a digital dynamic interface leads to an enormous increase in options about tools to use and accordingly options for the reader. The dynamic elements used in the current study had labels which informed the students about the use of the dynamic functions and the participants also received instruction before the test. Despite these instructions many participants still did not use the dynamic functions. Part of the explanation to the unwillingness to explore the digital environment can be due to unfamiliarity with the digital frame. As shown in previous research (see e.g., Dyrvold, 2022 and Ilovan et al., 2018) dynamic functions are relatively sparsely used in contemporary digital teaching materials. This means that even students who are used to working with digital textbooks during mathematics lessons, can be assumed to be inexperienced in using dynamic functions. Everyone who as a user has experienced a transition from one familiar digital platform to another, or the need to orient in a new digital administrative tool, likely recognise the frustration and lack of grit that may lead to abandoning to even try to grasp the new functions. When students choose to answer tasks on paper instead of digitally (Davis et al., 2021), the choice makes sense because the risk of misunderstanding how to present a solution using a pen is minimal, whereas missing a digital option is something that can happen. Bearing that in mind, the large share of students who missed the opportunity to use the highly dynamic and interactive elements type IV and V despite information just before the test situation, is less surprising.

As the use of computer-based assessments are spreading, we see a substantial risk for an increase in validity issues stemming from unfair demands of digital skills or of willingness to explore the media. If computer-based assessments are used as diagnosis tools or as part of formative assessments the risk is not as alarming because in a less stressful situation, it is more likely the test taker during the test develops a sensitivity to context and thereby manages to benefit from using the offered dynamic elements as expected. What we learn from this study is therefore applicable predominantly for assessments used for grading or to rank participants.

In retrospect, follow up interviews with the participants who did not use the dynamic elements would have contributed to the study. More participants in the *Show* group would also have strengthened the reliability of the results and there is a chance that more participants would have explored the content and eventually used the dynamic elements if the test was part of their mathematics course. These circumstances are important to address, and it is possible the results of a follow up study would differ to some extent. Despite these development areas however, the differences in use between more or less dynamic and interactive elements and between students who get the opportunity to try the dynamic elements beforehand or not, are convincing. Ideally the study would have been designed with matched pairs in the Try and Show group. The results from the survey to teachers and to the participants (Table 1 & 2) do however reveal that the participants in the two groups have fairly similar experiences of digital teaching materials and grades in mathematics and Swedish and it is therefore likely that the differences between the groups stem from differences in instruction.

5.1 Conclusion

Based on the statistics and the analysis according to the AOT perspective two main conclusions are drawn. Firstly, if dynamic elements utilised in a test are present in different contexts in several digital devices that students have used earlier, it is likely that these experiences can be transferred to the test situation. Because the layout of a test is likely to differ from students' previous experiences it is however suggested to at least show how the elements appear in the digital environment and make room for possible questions. Secondly, there is a substantial risk to overestimate students' capability to successful transfer from previous digital experiences and their capability to be sensitive to the new digital context (a test). If the capability to correctly use a dynamic element is not part of the construct being assessed, it is therefore recommended the students are given the opportunity to use all dynamic and interactive elements in a digital environment similar to the test before the test. The first conclusion is based on results where dynamic elements where click to open (type II) and film (type III) are used, and the second conclusion is based on results for the more dynamic and interactive elements (type IV-V). It is stressed that the options and buttons used in element type IV are very similar to response options used in online formulas or multiple-choice questions that are used very frequently in today's society. Despite that, many participants who did not try using the element in a similar context as in the test were not able to transfer previous experiences and thus missed opportunities in the test situation. This result highlights mode effect as a potential threat to the validity of an assessment, in particular when the demand for interaction is of a different kind than what is experienced by most citizens.

Acknowledgements

This work was supported by the Swedish Research Council [grant number 2019-05005].

References

Aldon, G. & Panero, M. (2020). Can digital technology change the way mathematics skills are assessed? *ZDM*, *52*(7), 1333–1348. https://doi.org/10.1007/s11858-020-01172-8

Baccaglini-Frank, A. (2021). To tell a story, you need a protagonist: How dynamic interactive mediators can fulfil this role and foster explorative participation to mathematical discourse. *Educational Studies in Mathematics, 106*(2), 291–312. https://doi.org/10.1007/s10649-020-10009-w

Barana, A., Marchisio, M., & Sacchet, M. (2021). Interactive feedback for learning mathematics in a digital learning environment. *Education Sciences, 11*(6), 279–290. https://doi.org/10.3390/educsci11060279

Bennett, R. E. (2015). The changing nature of educational assessment. *Review of Research in Education, 39*(1), 370–407. https://doi.org/10.3102/0091732X14554179

Bennett, R. E., Braswell, J., Oranje, A., Sandene, B., Kaplan, B., Yan, F. (2008b). Does it matter if I take my mathematics test on computer? A second empirical study of mode effects in NAEP. Journal of Technology, Learning, and Assessment, 6(9), 1–38.

Bennett, S., Maton, K., & Kervin, L. (2008a). The "digital natives" debate: A critical review of the evidence. *British Journal of Educational Technology*, *39*(5), 775–786. https://doi.org/10.1111/j.1467-8535.2007.00793.x

Cohen, J. W. (1983). *Statistical power analysis for the behavioral sciences* (2nd ed.) Lawrence Erlbaum Associates.

College Board. (2022). Assessment framework for the digital SAT suite, version 1.0 (June 2022). College Board.

Dadey, N., Lyons, S., & DePascale, C. (2018). The comparability of scores from different digital devices: A literature review and synthesis with recommendations for practice. *Applied Measurement in Education, 31*(1), 30–50. https://doi.org/10.1080/08957 347.2017.1391262

Davis, L. L., Morrison, K., Zhou-Yile Schnieders, J., & Marsh, B. (2021). Developing Authentic Digital Math Assessments. Journal of Applied Testing Technology, 22(1), 1–11. Retrieved from http://jattjournal.net/index.php/atp/article/view/155879

Dyrvold, A. (2022). Missed opportunities in digital teaching platforms: Under-use of interactive and dynamic elements. *Journal of Computers in Mathematics and Science Teaching*, 41(2), 135–161.

Gass, S. M., & Selinker, L. (1983). *Language transfer in language learning. Issues in second language research.* Newbury House Publishers, Inc.

Geraniou, E., & Jankvist, U.T. (2019). Towards a definition of "mathematical digital competency". *Educ Stud Math 102*, 29–45. https://doi.org/10.1007/s10649-019-09893-8

Goldstone, R., Son, J. Y., & Landy, D. (2008). A well grounded education: The role of perception in science and mathematics. Symbols and embodiment (pp. 327–356). Oxford University Press. http://jattjournal.net/index.php/atp/article/view/155879

Hamhuis, E., Glas, C., & Meelissen, M. (2020). Tablet assessment in primary education: are there performance differences between timss' paper-and-pencil test and tablet test among dutch grade-four students? *British Journal of Educational Technology, 51*(6), 2340–2358. https://doi.org/10.1111/bjet.12914

Hanho, J. (2014) A comparative study of scores on computer-based tests and paper-based tests. *Behaviour & Information Technology, (33)*4, 410–422. https://doi.org/10.1080/0144929X.2012.710647

Harris, D., Logan, T., & Lowrie, T. (2021). Unpacking mathematical-spatial relations: Problem-solving in static and interactive tasks. *Mathematics Education Research Journal*, *33*(3), 495–511. https://doi.org/10.1007/s13394-020-00316-z

Helsper, E. J., & Eynon, R. (2010). Digital natives: where is the evidence? *British Educational Research Journal*, *36*(3), 503–520. https://doi.org/10.1080/01411920902989227

Hoch, S., Reinhold, F., Werner, B., Richter-Gebert, J., & Reiss, K. (2018). Design and research potential of interactive textbooks: the case of fractions. *ZDM Mathematics Education, 50*(5), 839–848. https://doi.org/10.1007/s11858-018-0971-z

Ilovan, O.-R., Buzila, S.-R., Dulama, M. E., & Buzila, L. (2018). Study on the features of geography/sciences interactive multimedia learning activities (IMLA) in a digital textbook. *Romanian Review of Geographical Education, 7*(1), 20–30. https://doi.org/10.23741/RRGE120182

Junpeng, P., Krotha, J., Chanayota, K., Tang, K., & Wilson, M. (2019). Constructing progress maps of digital technology for diagnosing mathematical proficiency. Journal of Education and Learning, 8(6), 90–102. https://doi.org/10.5539/jel.v8n6p90

Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2013). The cost of concreteness: The effect of nonessential information on analogical transfer. Journal of Experimental Psychology. Applied, 19(1), 14–29. https://doi.org/10.1037/a0031931

Lemmo, A. (2021). A tool for comparing mathematics tasks from paper-based and digital environments. International Journal of Science and Mathematics Education, 19(8), 1655–1675. https://doi.org/10.1007/s10763-020-10119-0

Lobato, J. (2012). The actor-oriented transfer perspective and its contributions to educational research and practice. *Educational Psychologist, 47*(3), 232-247. https://doi.org/10.1080/00461520.2012.693353

Lobato, J. & Hohense, (2021). Current conceptualisations of the Transfer of learning and their use in STEM education research. In C. Hohensee & J. Lobato (Eds.), *Transfer of learning: Progressive perspectives for mathematics education and related fields* (pp. 3–26). Springer.

Lobato, J., & Siebert, D. (2002). Quantitative reasoning in a reconceived view of transfer. *The Journal of Mathematical Behavior, 21*(1), 87–116. https://doi.org/10.1016/S0732-3123(02)00105-0

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American psychologist*, *50*(9), 741–749.

Mullis, I. V. S., & Martin, M. O. (Eds.). (2017). *TIMSS 2019 Assessment Frameworks*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: http://timssandpirls.bc.edu/timss2019/frameworks

Nathan, M. J. & Alibali, M. W. (2021). An embodied theory of transfer of mathematical learning. In C. Hohensee & J. Lobato (Eds.), *Transfer of learning: Progressive perspectives for mathematics education and related fields* (pp. 27–58). Springer.

OECD (2013), PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy, OECD Publishing. http://dx.doi.org/10.1787/9789264190511-en

OECD (2021). *21st-Century Readers: Developing Literacy Skills in a Digital World*. PISA, OECD Publishing. http://dx.doi.org/10.1787/9789264190511-en

O'Halloran, K. L., Beezer, R. A., & Farmer, D. W. (2018). A new generation of mathematics textbook research and development. *ZDM Mathematics Education, 50*(5), 863–879. https://doi.org/10.1007/s11858-018-0959-8

Prensky, M. (2001). Digital Natives, Digital Immigrants. *On the Horizon, 9*(5), 1–6.

Regeringen (2017). Uppdrag att digitalisera de nationella proven. https://www.regeringen.se/4a80ac/contentassets/03dee5c5cdf244afa053e26cf654a8d8/uppdrag-att-digitalisera-de-nationella-proven-m.m..pdf

Ripley, M. (2009). Transformational computer-based testing. In F. Scheuermann & J. Björnsson (Eds.), *The transition to computer-based assessment* (pp. 92–98). Office for Official Publications of the European Communities

Smolinsky, L., Marx, B. D., Olafsson, G., & Ma, Y. A. (2020). Computer-based and paper-and-pencil tests: A study in calculus for STEM majors. Journal of Educational Computing Research, 58(7), 1256–1278. https://doi.org/10.1177/0735633120930235

Thorndike, E. L., & Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other functions. *Psychological Review, 8*, 247–261.

Usiskin, Z. (2018). Electronic vs. paper textbook presentations of the various aspects of mathematics. *ZDM Mathematics Education*, 50(5), 849–861. https://doi-org.ezproxy.its.uu.se/10.1007/s11858-018-0936-2

Yerushalmy, M., & Olsher, S. (2020). Online assessment of students' reasoning when solving example-eliciting tasks: Using conjunction and disjunction to increase the power of examples. *ZDM Mathematics Education, 52*(5), 1033–1049. https://doi.org/10.1007/s11858-020-01134-0

Research Article **LUMAT Special Issue 2023**

Comparing the integration of programming and computational thinking into Danish and Swedish elementary mathematics curriculum resources

Raimundo Elicer[1], Andreas Lindenskov Tamborg[1], Kajsa Bråting[2] and Cecilia Kilhamn[3]

[1] University of Copenhagen, Denmark
[2] Uppsala University, Sweden
[3] University of Gothenburg, Sweden

Computational thinking has become part of the mathematics curriculum in several countries. This has led recently available teaching resources to explicitly integrate computational thinking (CT). In this paper, we investigate and compare how curriculum resources developed in Denmark — digital teaching modules — and Sweden — printed mathematics textbooks — have incorporated CT in mathematics for grades 1–6 (age 7–12). Specifically, we identify and compare the CT and mathematical concepts, actions, and combinations in tasks within these resources. Our analysis reveals that Danish tasks are oriented toward CT concepts related to data, actions related to programming, and mathematical concepts within statistics. This is different from Swedish tasks, which are oriented toward CT concepts related to instructions and commands, actions related to following stepwise procedures, and mathematical concepts related to patterns. Moreover, what is most dominant in one country is almost or completely absent in the other. We conclude the paper by contrasting these two approaches with existing knowledge on computational thinking in school mathematics.

ARTICLE DETAILS

LUMAT Special Issue
Vol 11 No 3 (2023), 77–102

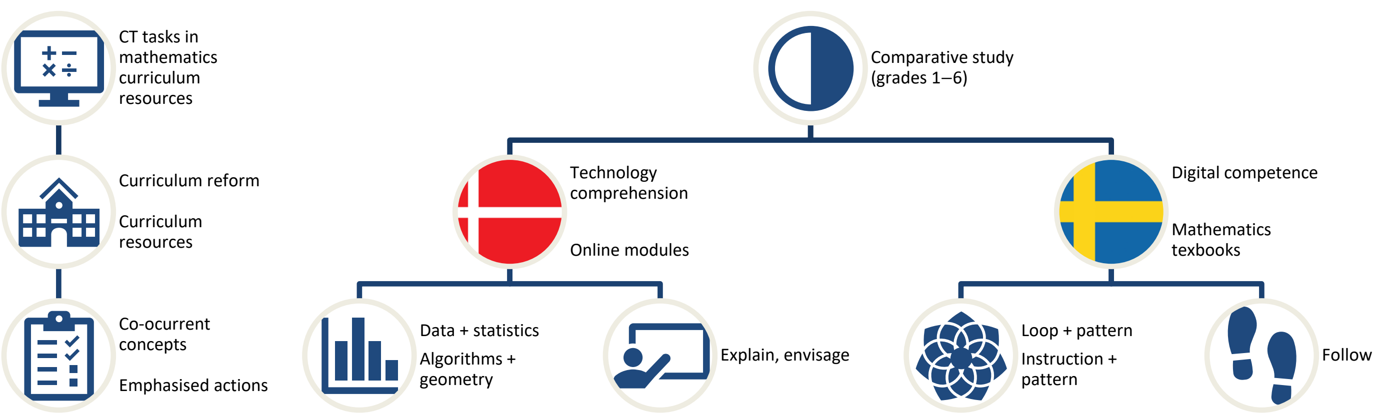
Received 15 February 2023
Accepted 13 November 2023
Published 24 November 2023

Pages: 26
References: 34

Correspondence:
reli@ind.ku.dk

https://doi.org/10.31129/LUMAT.11.3.1940

Keywords: computational thinking, curriculum resources, mathematics, programming



 | CC BY 4.0

1 Introduction

Computational thinking (CT) is a term coined in the scientific literature and first introduced by Papert (1980) as an educational construct. It roughly refers to a set of knowledge and skills necessary to frame problems in such a way a computer can carry out its solution without creating new problems (Wing, 2006). As a developing concept, it has a myriad of connotations that range from an attempted precise definition (Shute et al., 2017) to a collection of practices (Pérez, 2018; Weintrop et al., 2016). Some scholars argue for CT being more of a way of thinking, rather than being related to computing (Li et al., 2020), a conceptual foundation demonstrable 'with or without the assistance of computers' (Shute et al., 2017) and thus denoting programming as a separate skill. In general, CT is a highly ambiguous term. Palts and Pedaste (2020), for example, recently identified 65 distinct definitions of the term. In this article, we regard computational thinking as a set of constructs empirically observable in human productions particularly in school curricula[1]. CT constructs include programming knowledge and skills, data-handling practices, computational problem solving, modelling, algorithms and simulations. Rather than being guided by a specific definition of CT, this paper takes an outset as what is being referred to as such in Danish and Swedish curricula, which we unfold in the following subsections. It is worth noting that although Wing (2006) emphasized that computational thinking cannot be reduced to programming, many definitions include programming as a sub-element (Bocconi et al., 2022). Therefore, we consider programming to be sub-component of computational thinking.

Since Wing (2006) revived the concept, CT has come to be seen as a teachable competence beyond the domain of computer science, and curricula in many countries have expanded to include elements of CT (Bocconi et al., 2022). However, there are tremendous differences in the implementation strategies adopted by countries, and while some have established new subjects to address CT, e.g., Computing in England (Department for Education, 2013), other countries have revised existing subjects to include elements of CT, such as Sweden and France (Modeste, 2018). In the case of the latter, the CT elements have often been included in the mathematics curriculum.

---

[1] In this article, following Remillard (2005), we refer to the term curriculum as *formal* curriculum, namely the 'goals and activities outlined by school policies or designed in textbooks' (p. 213), distinct from the curriculum *intended* by teachers' aims and *enacted* in actual classroom practice.

The question of how to establish and exploit synergies between CT and mathematics education is a topic that has been subject to much research and unresolved discussion (see, e.g., Kohen-Vacs et al., 2020). An increasing amount of literature has proposed ways of framing how relationships between CT and mathematics are or should be (e.g., Benton et al., 2017; Gadanidis, 2017). This need for a better understanding has become increasingly important because a growing number of mathematics curriculum resources that include CT have begun to emerge. Yet, more attention has been given to developing ways in which mathematics and CT *could* or *should* be combined than to studying how they *in fact* are combined in available teaching resources (Tamborg et al., 2023). This study focuses on the Danish and Swedish cases, which are of particular interest as they are two neighbouring countries with shared traditions on the aims and approaches to school mathematics (Dahl & Stedøy, 2004). Yet, they have chosen two different ways of implementing CT on a policy level (Helenius & Misfeldt, 2021), whereby CT and mathematics competence descriptions and learning goals tend to be juxtaposed in the Danish case and integrated in the Swedish case (Tamborg et al., 2023). However, it is well known that curriculum policy does not correspond 1:1 to textbook material (Bråting & Kilhamn, 2022). The aim and contribution of this study is to investigate how CT and mathematics are represented and combined in textbook material in Denmark and Sweden and to engage in a discussion of the possible implications of such differences and/or similarities. We conduct this analysis as a comparative quantitative study of curriculum resources available in Denmark and Sweden by investigating what characterises them and how they differ.

We conduct this characterisation building from the suggested computational concepts (know-what) and practices (know-how) by Brennan and Resnick (2012). Although we expect the presence of CT and mathematical concepts in the curriculum resources, we aimed at pinpointing how they are represented and combined. Furthermore, based on Benton et al.'s (2017) framework for actions to embed CT in mathematics, we sought to identify the types of activities students should engage with. Thus, we operationalise our aim by addressing the following research question: *What are the combinations of CT and mathematical concepts and actions involved in Danish and Swedish curriculum resources, and how do they differ?*

We find this research question important and timely because CT is still a new component of mathematics education, which has become mandatory in many places and few teachers are trained in. Curriculum resources are likely to play a pivotal role in CT teaching in the mathematics classroom. Thus, mathematics teachers' experience of teaching CT is likely to be heavily rooted in the teaching resources that are available (Børne- og undervisnings ministeriet [BUVM], 2021b).

Since Denmark and Sweden are taking different strategies, we begin by briefly describing how CT is related to their mathematics curriculum and the type of curriculum resources that are available.

1.1 Programming as part of the mathematics curriculum in Sweden

In the work leading up to the revised K–9 national curriculum implemented in 2018, the Swedish National Agency of Education undertook the task of strengthening what was referred to as students' digital competence, described as an overarching competence area with no fixed content (Olofsson et al., 2021). The idea was that the content of the national curriculum should be successively renegotiated to include the digital competences that were relevant in the surrounding society. Consequently, this led to revisions of all major syllabi, in which the responsible use of digital media and its social, ethical, and legal aspects was categorised within social science and controlling objects by means of programming became part of the subject technology, while learning programming as such was integrated into mathematics throughout all grade levels (Heintz et al., 2017). In the mathematics syllabus, described in the national curriculum, programming was incorporated under the core content of algebra at all grade levels, described in the following ways[2] (Swedish National Agency of Education, 2018, pp. 56–59): for grades 1–3 (age 7–9): 'How unambiguous, step-by-step instructions can be constructed, described, and followed as a basis for programming. The use of symbols in step-by-step instructions.' For grades 4–6 (age 10–12): 'How algorithms can be created and used in programming. Programming in visual programming environments.' In grades 7–9 text-based programming is also added.

Helenius and Misfeldt (2021) emphasise that programming itself is in focus in the Swedish syllabi, and that the curriculum does not describe how programming can be used as a mathematical tool. Another characteristic of programming in the Swedish curriculum is that it primarily specifies a number of practices that students should be able to perform, whereas programming concepts are more or less absent.

In Sweden, teaching resources are primarily produced by private publishers, not by the National Agency of Education. Since 2018, several textbook producers have

2 This is the official English translation.

made efforts to include CT by developing teaching resources that address programming in mathematics. For grades 1–6, publishers have mainly done so by incorporating CT into the ordinary mathematics textbooks (Bråting & Kilhamn, 2022).

1.2 Technology comprehension in the Danish mathematics curriculum

Denmark has not yet made the final decision on revising the curriculum to include CT. In 2018, BUVM, however, launched a pilot project in which 46 schools across the country were to implement a new subject called Technology Comprehension (BUVM, 2018). The Danish approach sought to gain initial experiences with two different models of implementing technology comprehension to systematically research the effects of these approaches and, ultimately, inform a future, national-scale curriculum revision. The two strategies were 1) technology comprehension as a subject and 2) technology comprehension as an integrated part of existing subjects, such as Danish, mathematics, social sciences, science, physics/chemistry, craft and design, and the arts (BUVM, 2018). Both implementation strategies began with developing a curriculum for a subject in its own right. The individual components of this curriculum were then distributed among the subjects into which it should be integrated. The purpose declaration of technology comprehension emphasised a critical mindset and democratic values, reading 'students shall develop competencies and obtain skills and knowledge so that they constructively and critically can participate in the development of digital artefacts and understand their importance' (BUVM, 2019, authors' translation). This broad focus was also reflected in the four competency areas described: *Digital empowerment*, the critical exploration and analysis of how technology shapes our lives; *Digital design and digital design processes*, framing problems and generating solutions through iterative processes, which can lead to prototypes; *Computational thinking*, the ability to develop solutions to complex problems, the ability to make abstractions regarding phenomena and relationships, and computers' ability to process information; and *Technological agency*, the ability to understand and use digital technology to develop digital artefacts.

The two countries typify and classify aspects that can be considered part of CT differently. In Sweden, programming is the explicit manifestation of digital competence in the mathematics curriculum. In Denmark, CT is considered a competence area, while programming is left as a skill within technological agency. Therefore, we use CT as an inclusive term that gives justice to what programming entails in the Swedish curriculum and makes an explicit addition of programming into CT in Denmark.

In the following section, we briefly outline existing research on CT and mathematics to position the contribution of this paper in relation to the described body of knowledge.

2 Perspectives on relationships between mathematics and CT

During the past decade or so, an increasing body of knowledge on the relationship between CT and mathematics has emerged. This research follows several strands. One perspective has focused on unfolding the theoretical differences between mathematics and CT from an educational perspective and, in addition, developing arguments regarding how they could and why they should be connected. Along these lines, Pérez (2018) has argued that an essential difference between mathematics and CT is their orientations. In his view, mathematics tends to be inwards-oriented in the sense of being abstract and predominantly focusing on understanding disciplinary concepts and terms. In Pérez's (2018) view, CT is, on the contrary, more outward-oriented in that methods, concepts, and ways of thinking always are taught, learned, and applied in relation to practical problems in the real world. He argues that one powerful potential of integrating CT and mathematics is that it allows mathematics education to appeal to a broader and more diverse group of students, without favouring mathematics-advantaged students.

Gadanidis (2017) makes a similar proposition by arguing that, while mathematics and CT share a focus on logical structures and modelling, they operate within distinct epistemological frames. In his view, the frame of mathematics is associated with being a mathematician and engaging in mathematical practices, while the frame of computational thinking emphasises productive actions and their role in task optimisation. Gadanidis (2017) argues that CT can support mathematics education by, among other things, increasing students' agency, supporting abstraction, and enabling automation. Research however also found the overlap between mathematical and computing languages overlap to be a potential source of confusion. For example, Bråting and Kilhamn (2021) showed that symbols from the two domains can carry different meanings (e.g., the equal sign) and that different symbols carry the exact same meaning (e.g., modular arithmetic).

While these contributions indicate both good reasons to integrate CT into mathematics education and the potential pitfalls, they offer only a little advice regarding how to achieve such integrations. Weintrop and colleagues (2016) took this a step further and developed a taxonomy of four computational thinking practices for mathematics and science teaching. These practices are classified into data, modelling and simulation, computational problem-solving, and systems thinking (Weintrop et al., 2016). Each of these practices is described in terms of their taxonomical levels, from their most basic components to more advanced ones.

The work described above primarily develops arguments for why and how CT and mathematics *should* and *could* be combined. Fewer studies have conducted empirical descriptive work that investigates and compares what *in fact* is done in curriculum resources from different contexts, here among our recent work. Bråting and Kilhamn (2022) developed an analytical tool to analyse Swedish textbook tasks, adapting Brennan and Resnick's (2012) and Benton et al.'s (2017) frameworks for action. Analysing CT tasks designed for the Danish mathematics curriculum, Elicer and Tamborg (2022) took a grounded approach without any *a priori* defined categories by means of open, comparative, and iterative coding. A corollary to this study is that the categories can be approximated by identifying whether a task includes CT or mathematical concepts and actions and combinations thereof. We build on this work to further characterise and compare these teaching resources at a more systematic level.

3 Methodology

In what follows, we describe and justify the analytical strategy we will use to address the research question. Because the study takes available curriculum resources for elementary school (grades 1–6) as a point of departure, we first describe the sources of data and the selection process, leading up to specifying the unit of analysis. Second, we describe and argue for the analytical tool and its connections to existing theoretical frameworks. Finally, we display the strategies used to process and summarise the data analysis in light of the research question.

3.1 Data sources and selection of tasks

In Denmark, the subject Technology Comprehension has not yet become part of the mandatory curriculum. However, as part of the pilot project, expert groups developed a series of teaching modules oriented toward each of the two strategies for technology comprehension (as a subject in its own right and integrated into existing subjects), which are publicly available.[3] These modules include a declaration of competency areas and learning goals, an overarching scenario and problem statement, a sequence of tasks and resource banks. These resources were initially developed in 2019, and some of them were updated in 2021 as a result of the pilot project (BUVM, 2021b). We take these materials to be included in our analysis, since the resources constitute the only teaching materials that integrate technology comprehension in mathematics in compliance with the pilot curriculum.

In principle, there are 18 teaching modules developed for mathematics in grades 1 to 6, but four of them only cover technology comprehension areas related to digital empowerment, digital design and design processes, and sub-areas of technological agency outside of programming (Elicer & Tamborg, 2022). Therefore, a total of 14 teaching units include CT learning goals embedded into mathematics.

The modules are designed to be approached during several lessons, and they follow a general structure consisting of three phases: introduction, challenge and construction, and outro-phase (BUVM, 2021a). In turn, these phases are subdivided into tasks signposted with numerals and headings, sometimes subdivided into parts interpreted as separate tasks. We took these tasks as units of analysis in order to have fair ground of comparison to the Swedish curriculum material, which are mainly organized in smaller tasks. A total of 165 tasks were analysed in the Danish material.

In Sweden, since CT is an integral part of the mathematics curriculum, a fairly high volume of digital and printed resources are already available. In this study, we have chosen to restrict our analysis of the Swedish resources to printed mathematics textbooks because the programming content is then included in a well-known mathematical context and can be expected to be in use for a longer time as compared to digital resources, which are revised more often. We screened 56 Swedish mathematics textbooks for grades 1–6, of which 33 did not contain any tasks explicitly or implicitly labelled as belonging to CT (Bråting & Kilhamn, 2022). The resulting 23 textbooks included CT as chapters or sections titled ‘programming’ and ‘programming and pattern’, amounting to a total of 390 tasks, which are treated as units of analysis. The books belong to the following four series:

3 www.tekforsøget.dk/forlob

- Favorit matematik, grades 1–6, published by Studentlitteratur.
- Mondo matematik, grades 1–6, published by Gleerups.
- Singma matematik, grades 1–5, published by Natur & Kultur.
- Prima matematik, grades 1–6, published by Gleerups.

We explained the relevance and timeliness of comparing the curriculum resources from Denmark and Sweden. We acknowledge that the status of the curricular reforms and genre of their resources differ. However, the selection of tasks deals with these issues to a large extent. For both countries, the sources of data include relatively stable CT curriculum resources explicitly embedded in a mathematical context. All tasks are explicitly or implicitly related to CT, either by including CT concepts and actions or by appearing under a heading that relates them to CT. They are developed for grades 1–6 and contain a handleable number of signposted tasks, available for teachers to make use of in class without further instruction. Thus, we deem our data reasonable for comparison.

3.2 Theoretical underpinnings and analytical tool

Addressing our research question requires identifying CT and mathematical content in tasks that are included in Danish and Swedish mathematics curriculum resources, as well as how CT is combined with mathematics in these two contexts. In order to achieve this, we must identify the domain-specific aspects in the units of analysis. Based on the description of CT practices and concepts in the work of Brennan and Resnick (2012) and design principles for programming activities in mathematics developed by Benton et al. (2017), Bråting and Kilhamn (2022) constructed a framework suitable for analysing textbook tasks. In the following, we describe how we have applied these theoretical underpinnings to support the identification of the domain-specific *actions* and *concepts* involved in the selected tasks.

3.2.1 Concepts

In order to characterize CT resources for mathematics, we need to identify aspects of CT and aspects of mathematics that came to the fore in the tasks. Our analysis is partly based on identifying and distinguishing between two types of concepts, *CT concepts* and *mathematical concepts*. All of these are identified via the explicit and meaningful use of words in the context tasks. That is, we identify the occurrence of a signifier in the form of an explicit term, with the occurrence of the signified concept. Although a concept is much more than the words used to represent it (Wedman, 2020), in an analysis of written text, it is only through words, symbols and images that the intended concepts are available.

CT concepts are those that do not pertain to school mathematics in its traditional sense and belong to the task with the purpose of exploring and learning a computational idea (Li et al., 2020). Aside from programming concepts — e.g., code, algorithm, and condition — we consider concepts from CT, namely those related to (computer) modelling, data practices, and structures. This first step in identifying CT concepts is inspired by Brennan and Resnick's (2012) framework for CT, whose first dimension consists of a closed set of computational concepts: sequences, loops, parallelism, events, conditionals, operators, and data. However, while these predefinitions of CT concepts are helpful, they are not necessarily exhaustive. In our analysis, we follow a grounded approach in which the concepts described above serve as an important source of inspiration. The Danish teaching resources include a list of technological disciplinary concepts. For example, the task in Figure 1 comes from the module 'Concept of chance', which declares two competency areas from CT, namely programming and user studies and redesign. Therefore, *data* is not a CT concept in this context but, rather, a mathematical one. In Swedish textbooks and syllabus, computational concepts are referred to as programming concepts. For example, the concept *code* in the task displayed in Figure 2 is unambiguously a CT concept.

Mathematical concepts are those traditionally belonging to school mathematics, with an emphasis on the mathematical ideas to be learned in the particular context, including the mathematics sub-area and learning goals. For example, Figure 1 displays a Danish task that instructs students to play a dice game in pairs, register the results and winner of each play, and discuss whether the game is fair. Here, the mathematical (statistical) concepts highlighted are *fairness* and *data*. In the Swedish task shown in Figure 2, geometrical figures are represented in order to be identified. The one mathematical concept involved is *pattern*.

3.2.3. Subject loop

Students must play the game: "The difference between two dice". The game involves rolling two dice. The difference is determined and noted in a table. See *Table for Rolling Two Dice (Worksheet A1)*. Player A wins if the Difference is 2 or 4.

Students must gain experience with the game by playing together two by two with ordinary dice (one blue and one red). A student is defined to be player A and one is defined to be player B.

Dice 1	Dice 2	Difference	A wins
5	1	4	1
3	3	0	0

Figure 3 Table for Rolling Two Dice – Difference

If A wins, 1 is written in the last column. If A loses, 0 is written, so that 1 indicates a positive answer and 0 a negative outcome.

Students get 15 min. to play the game, after which they in the group should discuss whether the game is fair. They have to argue for their answers based on their data.

Students' experiences from the game are gathered in class and reasons why the game is unfair are shared.

Figure 1. Task belonging to the Danish course ‘Concept of chance’

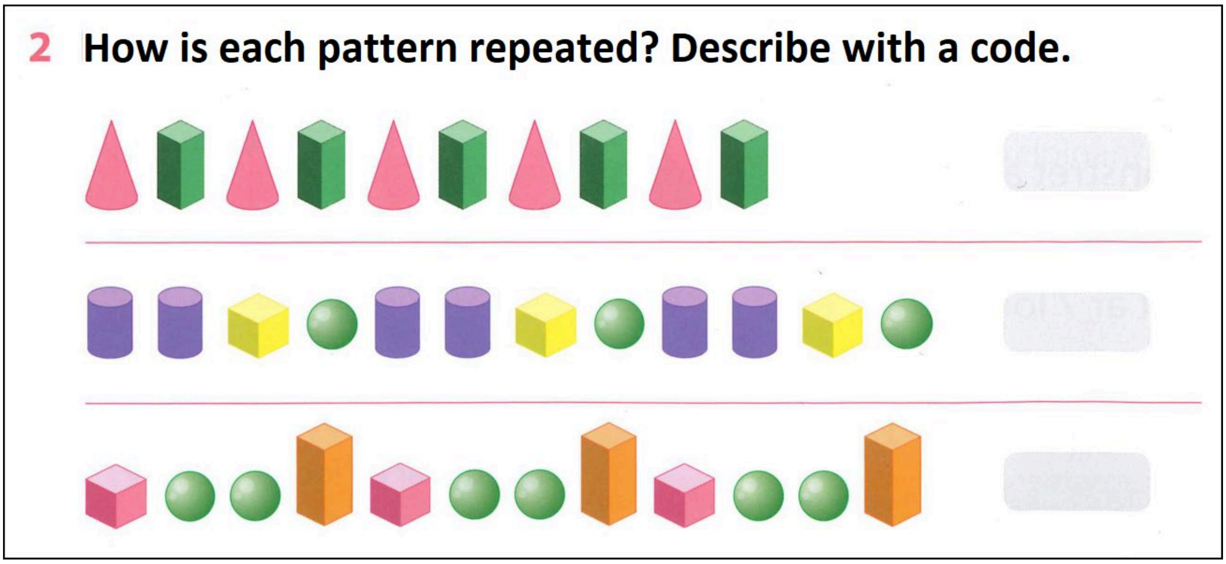


Figure 2. Programming task from Swedish textbook Singma 3B

3.2.2 Actions

In addition to the invoked concepts, we classify the CT-related activities students are explicit asked or expected to engage in when doing the task as *actions*. Though Brennan and Resnick (2012) defined *computational practices* as their second dimension in framing CT, these practices do not highlight CT’s relationship to mathematical learning. As part of the *ScratchMaths* project, to answer the question of what programming in Scratch can do for students’ mathematical learning, Benton et al. (2017) delineated five such activities. These five activities formed what they refer to as the 5E pedagogical framework. One caveat regarding this framework is that it was developed to *design* programming activities in Scratch, as opposed to *analysing* tasks at face value. In other words, it is a framework for action, which includes 'prescriptions for pedagogical strategies' (diSessa & Cobb, 2004, p. 81) and can provide effective heuristics for designing and teaching. We have adapted the action framework of Benton et al. (2017) into six actions suitable for analysing the Swedish data. Detailed examples of each action can be found in Bråting and Kilhamn (2022). The actions are:

a) Follow, i.e., follow a procedure, follow stepwise instructions, or repeat or continue a pattern.
b) Figure out, i.e., work out a procedure, a rule, or a pattern
c) Debug, i.e., find mistakes in a pattern or debug a code
d) Program, i.e., form and create, give instructions, create a pattern, write code, or represent with symbols.
e) Explain, i.e., using words/natural language to explain or describe a procedure, a rule, a pattern, or a concept.
f) Envisage, i.e., predict what will happen or reflect on potential outcomes when conditions or values are changed.

For example, in Figure 1, students should *follow* (a) instructions to play a game and register its outcomes. Although the task includes a discussion and plenum about the game's fairness based on its results, it is not the instructions for the game that must be figured out (b) or explained (e). The task in Figure 2 is coded as *figure out* (b) because students should work out the pattern that the figures follow. Although the concept of *code* is present, the task only asks them to describe the pattern with a code, not to create an original program or pattern (d).

The analytical tool described above could be seen as a compromise between the open, face-value coding of concepts and predefined practice-oriented categories. It has previously been successfully used to analyse Swedish textbooks (Bråting & Kilhamn, 2022). For comparison's sake, this is the analytical tool of choice to provide an initial characterisation and overview of tasks in the Swedish and Danish resources. However, addressing our research question requires us to further process these findings in the analysis.

3.3 Data processing and analysis

As stated, our research question focuses on investigating the characteristics of the concepts and actions included in the Danish and Swedish tasks, as well as how these were combined with mathematical content in these two contexts. In order to address this question, we began by coding the tasks to identify what concepts from mathematics and CT the tasks address and what types of programming actions they include. Our coding of concepts from mathematics and CT was based on the concepts explicitly mentioned in the tasks (e.g., triangle, area, addition, etc. and algorithm, loop, and bug). Next, we coded each task according to what type of actions it included. We conducted all coding manually in a spreadsheet. For each country's case, two researchers coded the tasks — concepts and actions — following three stages: 1) joint coding, 2) parallel coding with a later settling of eventual disagreements, and 3) the separate coding of the remainder. We remained in Stages 1 and 2 for approximately one-third of the total data for each country. This process left us with a descriptive account of the CT and mathematical content addressed in the tasks. After coding the material, we conducted four aggregations or summaries, which we report in the results section.

The first two aggregations result after counting the number of tasks that contain CT and mathematical concepts for each country's material. For comparison's sake, we report the percentage relative to the total number of tasks from each country; 165 from the Danish material and 390 from the Swedish material. The most frequent CT concepts are reported and compared in Table 1. Mathematical concepts are more diverse, and thus, we classified them into arithmetical (operations, number systems), geometrical (polygons, coordinate plane, angles) and statistical (probability, data) concepts. Given the strong emphasis on patterns and number sequences in Sweden, we designated these as a separate category. Mathematical concepts are reported in Table 2.

The next two data processing steps address combinations. First, we compare the co-occurrence of CT and mathematical concepts. For this, we count the number of tasks containing combinations of the CT and mathematical concepts reported in Tables 2 and 3. For example, the task in Figure 1 only refers to data in its mathematical (statistical) sense, namely as registered instances of a random process. Therefore, that task does not report such a co-occurrence. However, the task in Figure 2 includes a mathematical and a CT concept, *pattern* and *code*, respectively, so it is counted as a co-occurrence. This summary and comparison are reported in Table 3.

Finally, we compare the distribution of actions involved in tasks that include mathematical concepts. After filtering out those tasks that do not include mathematical concepts, we count the number of tasks, for each country, that involve each of the six actions of our analytical tool. These are reported in Figure 3 as percentages, for the sake of a fair comparison. It is important to note that tasks may include more than one CT or mathematical concept, as well as several actions. For this reason, the total does not always add up to 100%. We will return to this matter in the discussion. Below, we describe the results of our analysis described above.

4 Results

4.1 CT concepts

As described above, the first analysis concerns the number of tasks containing different CT concepts. Table 1 below represents the results of this and the relative distribution of CT concepts in the resources from the two countries.

Table 1. Overview of CT concepts and their distribution in Danish and Swedish tasks, absolute numbers and relative percentages.

	Danish tasks		**Swedish tasks**	
Instruction, command	12	7%	189	48%
Algorithm	19	12%	74	19%
Loop, iteration, repetition	0	0%	101	26%
Rule	1	1%	51	13%
Code	14	8%	59	16%
Condition	0	0%	37	9%
Bug, debugging	1	1%	17	4%
Data	20	12%	2	1%
Program, programming	33	20%	10	3%
Totals	**165**		**390**	

In Table 1, we see that the CT concepts of programming, algorithm, and data are the most frequent in the Danish tasks. The least represented CT concepts in Danish tasks are loop/iteration/repetition and condition, none of which are present at all. Rule and bug/debugging each appear in 1% of the tasks. As the reader may notice, the sum of the percentages in the Danish column does not add up to 100. This is because the percentages are computed against the total number of tasks, some of which may include more than one CT concept. However, 63 (38%) Danish tasks do not include CT concepts at all.

In contrast, only 34 (9%) of the Swedish tasks analysed include no CT concepts. In these materials, the most frequently occurring CT concepts are instruction/command, loop/iteration/repetition, and code. Instruction/command stands out by appearing in nearly half of the Swedish tasks. Approximately one quarter of the Swedish tasks include the concept of loop/iteration/repetition (the three terms are in here seen as representing the same concept). The least represented CT concepts are data, program/programming, and bug/debugging. The low representation of data stands out when compared to the Danish case, where it was the second most frequent concept. Moreover, program/programming is the second least represented CT concept in Sweden, while in Denmark, it was the most frequent CT concept.

4.2 Mathematical concepts

The second analysis concerns the mathematics concepts in the tasks. The result of this analysis is summarised in Table 2, which also shows the relative distribution of the mathematical concepts.

Table 2. Overview of mathematical concepts and their distribution in Danish and Swedish tasks in absolute numbers and relative percentages.

	Danish tasks		**Swedish tasks**	
Pattern, sequence	0	0%	131	34%
Geometrical concepts	56	34%	61	16%
Arithmetical concepts	17	10%	26	7%
Statistical concepts	35	21%	2	1%
Totals	**165**		**390**	

In Table 2, we see that patterns and sequences by far are the most frequently occurring mathematical concepts in the Swedish tasks and that geometrical concepts are the second most frequent.

Regarding the Danish tasks, Table 2 shows that geometrical concepts are the most frequent and that statistical concepts are the second most frequent. Patterns and sequences are absent, although it could be argued that several of the Danish tasks address patterns implicitly. One example of this is the task entitled ‘Design the class’s new clock’, in which students are to develop pattern-like figures in GeoGebra, which they can use as the background for a watch. The tasks, however, focus on design processes and do not deal with patterns in the mathematical sense of the word.

It is worth mentioning that, in both cases, the share of tasks that do not engage with mathematical concepts is similar. 150 (38%) of the analysed Swedish tasks and 71 (43%) of the analysed Danish tasks do not contain any explicit terms referring to mathematical concepts.

4.3 Co-occurrence of concepts

Overall, 44 (27%) of the Danish tasks contain both CT and mathematical concepts, as compared to 195 (50)% of the Swedish tasks. This feature is reflected in the structure and pace of the teaching resources. Several tasks in the Danish teaching modules are meant to focus on a particular concept, either mathematical or CT, leaving the connections between disciplines to a later wrap-up task (Elicer & Tamborg, 2022).

In relative terms, 53% of the Danish tasks that include CT concepts also contain mathematical concepts compared to 63% of the Swedish tasks. In other words, when tasks introduce or draw on CT concepts, they are, to a larger extent in Sweden than in Denmark, combined with mathematical concepts within them.

In what follows, we can see how these co-occurrences appear to be disaggregated by the type of CT and mathematical concepts.

Table 3. Co-occurrence of mathematical and CT concepts in both countries. Percentages are relative to the total number of tasks from each country.

	Pattern, sequence		**Geometrical concepts**		**Arithmetical concepts**		**Statistical concepts**	
Country	DK	SE	DK	SE	DK	SE	DK	SE
Instruction, command	0%	12%	3%	7%	1%	3%	1%	0%
Algorithm	0%	3%	7%	8%	1%	2%	0%	0%
Loop, iteration, repetition	0%	18%	0%	4%	0%	0%	0%	0%
Rule	0%	12%	1%	0%	0%	0%	0%	0%
Code	0%	4%	4%	1%	1%	2%	1%	0%
Condition	0%	0%	0%	2%	0%	0%	0%	1%
Bug, debugging	0%	3%	0%	0%	0%	0%	0%	0%
Data	0%	0%	2%	0%	2%	0%	9%	0%
Program, programming	0%	0%	5%	2%	2%	0%	3%	0%

The mathematical and CT concepts that most often co-occur in Denmark are data + statistics, algorithm + geometry, and program + geometry. The CT concepts of loop, condition, and bug and the mathematical concept of pattern are absent from the Danish tasks. The most frequently co-occurring mathematical and CT concepts in the Swedish data are loop + pattern, instruction + pattern, and rule + pattern. The only CT concept that is absent from the Swedish tasks is data. An observation that stands out from a comparative perspective is that the most frequent co-occurrence in Denmark (data + statistics) is completely absent in the Swedish material. Likewise, the three most frequent co-occurrences in Sweden are completely absent in the Danish material.

4.4 Actions

Figure 3 below illustrates the distribution of actions in the tasks from the two concerned countries. Each task can include more than one action.

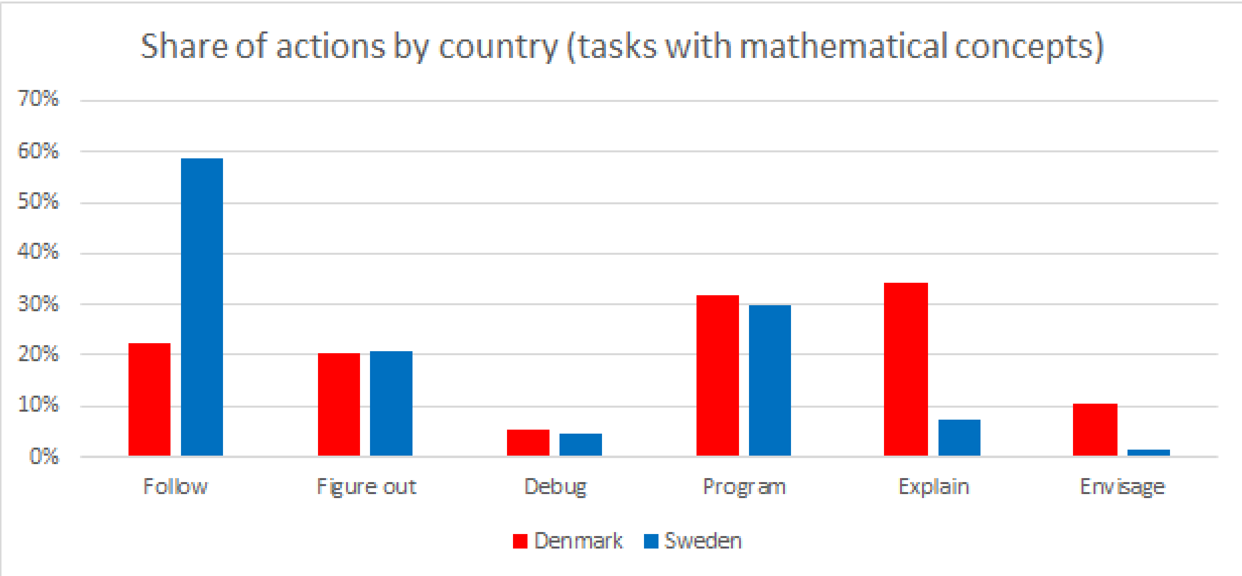


Figure 3. Distribution of the six different actions. Percentages are computed relative to each country's total number of tasks with mathematical concepts: 94 Danish tasks and 250 Swedish tasks.

The first thing to notice from Figure 3 is the difference in the share of tasks that include *following* a procedure and *explaining* it. The follow action is more than three times as frequent in the Swedish tasks as compared to the Danish tasks. With regard to tasks that include the action *explain*, we see that the Danish tasks include this action more than four times as often as the Swedish tasks. The reasons for this could potentially both be found in the format of the Danish tasks and in the content of the Danish technology comprehension curriculum. The template for the Danish teaching modules consistently includes time allocated to 'setting the scene' and 'wrapping up'. Activities in these sections predominantly consist of open questions for students and teachers to address, either in groups or, more often, in plenary classroom discussions. Particularly in the wrap-up sections, students are often asked to present and explain the outcome or process in which they were engaged. The high proportion of the action *explain* in the Danish tasks is also consistent with the emphasis given to democratic participation in digital contexts in the curriculum. For example, the goal of the technology comprehension content area modelling specifies that after grade 3, the student must 'be able to describe the reality represented by a model[4].' Addressing this goal is likely to include the use of natural language to describe a procedure, rule, pattern, or concept, which is how the action *explain* is defined in the analytical tool.

5 Discussion

In what follows, we discuss the findings presented in the results section in light of our research question and earlier research within the field.

5.1 Insights on concepts, data sources and analytical tool

Our results reveal that there are notable differences between what CT concepts are included in tasks designed for elementary school mathematics in the two countries and their relative distribution. To some extent, these differences can problematise the choice of data sources and analytical tool.

A significant share (38%) of Danish tasks do not include CT concepts, compared to a 9% of Swedish tasks. Such a sharp difference is consistent with the different way our data sources embedded CT into mathematics. On the one hand, the Danish resources cover combinations of mathematical, CT and other technology comprehension competency areas throughout the modules, which are declared on the front-matter of each material. For this reason, they do not necessarily cover exclusively CT areas nor do so throughout all the tasks in them (BUVM, 2021a). On the other hand, the selected Swedish mathematics textbooks included CT allocated as specific chapters or sections (Bråting & Kilhamn, 2022). It makes sense that a majority of the tasks within them are much more focused on CT.

The absence of mathematical concepts is significant in both cases, respectively 43% of Danish tasks and 38% of Swedish tasks. This feature reflects back on the coding of only explicit terms, as we argued for in section 3.2.1. Although both sources of data come from explicitly mathematics curriculum resources, our decision to code only for explicit terms in the tasks can hide many concepts implicitly involved in their

[4] https://emu.dk/sites/default/files/2019-02/7568_STIL_M%C3%A5l_Matematik_web_FINAL-a.pdf

interpretation and solution. However, the fact that the share of tasks with no mathematical concepts is fairly similar makes the comparisons in sections 4.3 and 4.4 more grounded. The coding of explicit terms may also explain why the CT concept *program/programming* is prominent among Danish tasks in comparison to *code*, while the opposite is true with Swedish tasks (Table 1). One could argue that these terms are interchangeable in one or both contexts, but we deem these terms to represent different concepts. A program (e.g., a function or algorithm) and the act of programming (closer to generic modelling) can be done with different codes and languages (Caeli & Yadav, 2020). More importantly, their synonymity would be an assumption we cannot make strictly based on the data. A limitation of this study's basis for comparison is the fact that, despite being updated after the pilot study, curriculum resources in Denmark are not yet massively used. Therefore, one necessary extension of this work includes analysing Danish textbooks once the reform rolls out, so that the massive use, stable status and number of tasks are more fairly comparable.

5.2 Co-occurrence and the significance of syllabi

We also see notable differences in how actions and CT concepts are combined with mathematical concepts. In the Swedish tasks, actions are highly skewed toward *following* a procedure (Figure 3) while following stepwise *instructions*, along with the concepts of *loop*, *iteration*, and *repetition*, as the most frequent CT concepts (Table 1). These CT concepts and actions are most frequently combined with mathematical content related to *patterns*. The Danish tasks, on the other hand, most frequently include the actions *explain* and *program* (Figure 3) and CT concepts related to programming and data (Table 1), which most frequently co-occur with statistical concepts (Table 3).

To some extent, these characterisations resemble the curricular decisions made reforming the mathematics curriculum to include CT in the two countries. In the Swedish mathematics syllabus described in the national curriculum, following stepwise instructions is explicitly mentioned in grades 1–3, and for grades 4–6, students must be able to program and use algorithms in visual programming environments. This is reflected in our results, in which loop, iteration, and repetition are all essential components of developing algorithms[5]. Based on the high frequency of mathematical concepts related to patterns in the Swedish data, we can speculate that decisions made at

[5] See, for example, https://www.bbc.co.uk/bitesize/guides/zg46tfr/revision/1

the national curricular level influences paths taken at the level of textbook designers and, consequently, perhaps also in the classroom. When CT was incorporated into mathematics in the Swedish curriculum, a choice had to be made. Either a new subject matter area could have been added, which would have given CT the same status as the traditional subject matter areas, or CT could be inserted into one of the existing subject matter areas. For some reason, algebra was pinpointed as the best place for CT in grades 1–6, not statistics or geometry, which are the most frequent content areas in the Danish tasks. Hence, in the Swedish resources, programming became tightly connected to patterns, a topic already present in algebra.

Similarly, we see a relationship between the Danish tasks and the technology comprehension syllabus. This syllabus has a strong emphasis on students' ability to, e.g., critically explore how technology shapes our lives, frame problems, and use digital technology to develop digital artefacts, all of which are integral to the four technology comprehension competency areas (BUVM, 2019). Such curriculum aims seem to align well with tasks in which students are to program, explain what they have accomplished, and combine data and statistics to inquire into societal phenomena from a mathematical and CT perspective.

This alignment between tasks and the mathematics curriculum in the two countries suggests that the curriculum documents are indeed significant for how curriculum resource developers have engaged in the integration of CT. This may seem obvious, but the alignment has significant implications. It points to important limitations of previous theoretical work on potential ways of establishing synergetic relationships between CT and mathematics. The taxonomy of CT practices in the mathematics classroom developed by Weintrop et al. (2016) and Pérez's (2018) arguments for the usefulness of an orientation toward real-world problems in CT for mathematics education are, no doubt, important contributions. They have broadened our understanding of the fundamental differences between mathematics education and CT and provided tools for navigating this new landscape. However, the results of this study remind us that curriculum resources are often developed to comply with curriculum policy, and the implementation of research into curriculum guidelines and resources may be influenced by a diversity of political factors (Aguilar & Castaneda, 2022). In such contexts, theoretical ideas and suggestions regarding synergetic relations between mathematics and CT are likely to be thought of as useful only to the extent that they align with decisions in the mathematics syllabus. Despite the fact that, for instance, data practices can theoretically constitute an obvious boundary object for mathematics and CT (Weintrop et al., 2016; Gould, 2021), this is of little to no relevance in the Swedish mathematics syllabus for elementary grades where programming is a mandatory part of algebra. If the analysis had been extended to resources for grades 7–9, the results may have been different because the Swedish curriculum does include the 'assessment of risk and chance based on computer simulations and statistical material' in the area of probability and statistics for these grades.

The mathematics curricula for elementary grades in Sweden and Denmark point to two ways in which curriculum policy may constrain the integration of CT into teaching resources, namely content itself and the level of specificity in relation to mathematical content. The two contexts differ in that the CT elements in the Danish mathematics curriculum are rather generic and not content-specific, while in the Swedish mathematics curriculum, CT components are more technically narrowed, leaving other issues to other subjects and educational levels (Helenius & Misfeldt, 2021). This could help explain the perhaps most outstanding result of our study: the most frequently co-occurring CT and mathematical concepts in the Danish resources (data + statistics), completely absent in the Swedish resources, are not as dominant compared to other frequent co-occurrences (Table 3). At the same time, the most frequently co-occurring CT and mathematical concepts in the Swedish resources (instruction/loop/rule + pattern), completely absent in the Danish case, standout drastically from other combinations (Table 3).

5.3 Bridging and resource structure

As our results showed, the available Danish resources tend to have a more even share of tasks involving mathematical concepts (Table 2). However, Swedish textbooks display a larger overall share of tasks that combine CT and mathematical concepts than Danish resources (Table 3). In fact, according to Figure 3, the same can be said regarding actions when mathematical concepts are involved. Although Danish teaching modules always included tasks combining concepts from both domains of knowledge, there were multiple instances in which smaller sub-tasks only included concepts from either mathematics or CT. These dissimilarities between Danish and Swedish tasks in the way they combine CT and mathematics can also be explained in terms of how they are situated in the structure of teaching resources, beyond the data sources we discussed in section 5.1.

At a structural level, the Danish tasks are organised as teaching sequences divided into several smaller steps, slowly progressing toward an end-goal. For example, the task in Figure 1 — with only mathematical (statistical) concepts — is one of many introductory activities about the notion of chance in the module. Later, the idea of the fairness of a game is connected to the difficulty of a game that students should program and adjust. This is different from the Swedish tasks, which are smaller and more independent. They are not, as such, embedded in the context of a larger inquiry. The task in Figure 2 — with co-occurrent mathematical and CT concepts — is self-sufficient and not at all interdependent with those in the same textbook chapter. In that sense, the results may reflect two distinct approaches in terms of connecting mathematics and CT, either *between* tasks or *within* tasks.

In our use of the term actions, we refer to CT actions in the context of mathematics teaching resources. We have thus not distinguished between mathematical and CT actions. Identifying such connections between domains within a task could be labelled as a CT and mathematical action simultaneously. In their study, Bråting and Kilhamn (2022) acknowledge that the analytical tool does not account for some tasks in which the actions' fields of origin are ambiguous. Benton and colleagues (2017) call this action 'bridging' computational and mathematical ideas. Some Danish tasks also pertain to this overarching category, which Elicer and Tamborg (2022) call 'operational integration'. In a few words, they represent a missing category of tasks in which opportunities to link mathematical and computational concepts become clearer. In a sense, such tasks seem to justify the integration of CT and mathematics by necessitating concepts from both fields within the same action.

6 Conclusion

Our research question aims to characterise the analysed resources by identifying CT and mathematical concepts, combinations thereof and actions involved in the tasks.

In terms of disciplinary concepts, at the K–6 level, resources illustrate a contrast between a notion of CT that focuses mostly on programming and algorithms and another that takes a broader view, one including the handling of data and computer modelling. Danish tasks involve a relatively even distribution of mathematical — arithmetical, geometrical, and statistical — concepts, while Swedish tasks are highly skewed toward patterns and sequences, which are absent in the Danish tasks.

As for combinations of mathematical concepts with CT concepts and actions, what is roughly most available in one is absent in the other. Danish tasks rely on the interplay between data as a CT and a statistical concept in its mathematics curriculum resources. This aspect is in line with the recent trend on developing a common data-scientific literacy as a merge between statistical, mathematical, and computational thinking (Gould, 2021). In turn, Swedish tasks focus on the combination of stepwise instructions, patterns, and number sequences; a strong focus on programming and early algebra as per the first wave of CT (Clements & Sarama, 1997).

In the introduction, we argued for the need of studies of curriculum resources that not only consider how CT and mathematics *could* and *should* be integrated, but how the two subjects *are actually* integrated in available resources. Given the novelty of CT's integration into mainstream curricula, tasks are what is available for teachers to implement the potential synergies this innovation may bring and, in turn, for students to make use of. On that line, our study provides two main contributions.

First, we have documented that two neighbouring countries with an otherwise shared tradition of mathematics teaching integrate CT into mathematics quite differently; a contrast that permeates its curriculum resources. From a comparative point of view, these empirical findings can inform future discussions about shortcomings and potentials in both approaches when policy decisions are translated into teaching materials. Therefore, our results can be sources of awareness and inspiration in the search for alternative curricular strategies than the ones currently adopted to integrate CT into mathematics.

Second, we constructed an analytical tool to conduct the first contribution building on the state of the art. Despite its acknowledged limitations, it has proven applicable in two highly different contexts with regards to implementation strategies and status, and types of curriculum resources. Considering that the interest in CT is a global trend, we envision our analytical tool as suitable to characterise and position curriculum resources from other contexts in relation to one another. The main focus of this paper has been devoted to curriculum material, thereby leaving out teachers' enactment of the resources in the mathematics classroom. This delineation implies constraints in terms of what this study can offer as insights on the implications for classroom practices. As argued previously, the novelty of CT in K-9 and teachers, limited knowledge of its associated concepts is likely to mean that they will largely lean on curriculum material. However, there remains an empirical question and, as more resources on CT and mathematics will continue to emerge, it is an important direction for future research to study how teachers choose and enact CT curriculum material in the classroom and with what learning outcomes for students.

References

Aguilar, M. S., & Castaneda, A. (2022). Out of the public eye: Researching political factors that influence the implementation of research knowledge as part of educational reforms and mathematics textbooks. *Implementation and Replication Studies in Mathematics Education*, *2*(1), 107–129. https://doi.org/10.1163/26670127-bja10001

Benton, L., Hoyles, C., Kalas, I., & Noss, R. (2017). Bridging primary programming and mathematics: Some findings of design research in England. *Digital Experiences in Mathematics Education, 3*(2), 115–138. https://doi.org/10.1007/s40751-017-0028-x

Bocconi, S., Chioccariello, A., Kampylis, P., Wastiau, P., Engelhardt, K., Earp, J., Horvath, M., Malagoli, C., Cachia, R., Giannoutsou, N., & Punie, Y. (2022). *Reviewing computational thinking in compulsory education: State of play and practices from computing education* (I. dos Santos, R. Cachia, N. Giannoutsou, & Y. Punie (Eds.)). Publications Office of the European Union. https://doi.org/https://doi.org/10.2760/126955

Bråting, K., & Kilhamn, C. (2021). Exploring the intersection of algebraic and computational thinking. *Mathematical Thinking and Learning*, *23*(2), 170–185. https://doi.org/10.1080/10986065.2020.1779012

Bråting, K., & Kilhamn, C. (2022). The integration of programming in Swedish school mathematics: Investigating elementary mathematics textbooks. *Scandinavian Journal of Educational Research*, *66*(4), 594–609. http://doi.org/10.1080/00313831.2021.1897879

Brennan, K., & Resnick, M. (2012). New frameworks for studying and assessing the development of computational thinking. In *Proceedings of the 2012 Annual Meeting of the American Educational Research Association* (pp. 1–25). American Educational Research Association.

Børne- og Undervisningsministeriet (2018). Handlingsplan for teknologi i undervisningen. https://www.uvm.dk/publikationer/folkeskolen/2018-handlingsplan-for-teknologi-i-undervisningen

Børne og Undervisningsministeriet (2019). *Fælles mål for teknologiforståelse* [Common goals for technology comprehension]. https://emu.dk/sites/default/files/2019-02/GSK.%20F%C3%A6lles%20M%C3%A5l.%20Tilg%C3%A6ngelig.%20Teknologiforst%C3%A5else.pdf

Børne- og Undervisningsministeriet (2021a). *Didaktiske prototyper - Format og vejledning* [Didactical prototypes – Format and guidelines]. https://tekforsøget.dk/wp-content/uploads/2021/06/Format-og-vejledning-til-didaktiske-prototyper-maj-2021.pdf

Børne- og Undervisningsministeriet (2021b). *Forsøg med teknologiforståelse i folkeskolens obligatoriske undervisning: Slutevaluering* [Experiment with technology comprehension in compulsory education: Final evaluation]. https://www.uvm.dk/-/media/filer/uvm/aktuelt/pdf21/okt/211004-slutevaluering-teknologoforstaaelse.pdf

Caeli, E. N., & Yadav, A. (2020). *TechTrends*, *64*(1), 29–36. https://doi.org/10.1007/s11528-019-00410-5

Clements, D. H., & Sarama, J. (1997). Research on Logo: A decade of progress. *Computers in the Schools*, *14*(1–2), 9–46. https://doi.org/10.1300/J025v14n01_02

Dahl, B., & Stedøy, I. M. (2004). A Nordic community: Ideas of education and democracy in mathematics. In I. M. Stedøy (Ed.), Mathematics Education - The Nordic Way (pp. 1–10). TAPIR.

Department for Education. (2013). National curriculum in England: Computing programmes of study. National Curriculum in England. https://www.gov.uk/government/publications/national-curriculum-in-england-computing-programmes-of-study/national-curriculum-in-england-computing-programmes-of-study diSessa, A. A., & Cobb, P. (2004). Ontological innovation and the role of theory in design experiments. Journal of the Learning Sciences, 13(1), 77–103. https://doi.org/10.1207/s15327809jls1301_4

Elicer, R., & Tamborg, A. L. (2022). Nature of the relations between programming and computational thinking and mathematics in Danish teaching resources. In U. T. Jankvist, R. Elicer, A. Clark-Wilson, H.-G. Weigand, & M. Thomsen (Eds.), *Proceedings of the 15th International Conference on Technology in Mathematics Teaching* (pp. 45–52). Aarhus University.

Gadanidis, G. (2017). Five affordances of computational thinking to support elementary mathematics education. *Journal of Computers in Mathematics and Science Teaching*, *36*(2), 143–151.

Gould, R. (2021). Towards data scientific thinking. *Teaching Statistics*, *43*(S1), S11–S22. https://doi.org/10.1111/test.12267

Heintz, F., Mannila, L., Nordén, L. Å., Parnes, P., & Regnell, B. (2017). Introducing programming and digital competence in Swedish K-9 education. In V. Diagené & A. Hellas (Eds.), *International Conference on Informatics in Schools: Situation, Evolution, and Perspectives* (pp. 117–128). Springer. https://doi.org/10.1007/978-3-319-71483-7_10

Helenius, O., & Misfeldt, M. (2021). Programmeringens väg in i skolan: en jämförelse mellan Danmark och Sverige. [Programming's way into school – a comparison between Denmark and Sweden]. In K. Bråting, C. Kilhamn, & L. Rolandsson (Eds.), *Programmering i skolmatematiken: möjligheter och utmaningar* (pp. 39–56). Studentlitteratur.

Kohen-Vacs, D., Kynigos, C., & Milrad, M. (2020). On the integration of learning mathematics and programming. In S.-C. Kong, H. U. Hoppe, T.-C. Hsu, R.-H. Huang, B.-C. Kuo, R. K.-Y. Li, C.-K. Looi, M. Milrad, J.-L. Shih, K.-F. Sin, K.-S. Song, M. Specht, F. Sullivan, & J. Vahrenhold (Eds.), *Proceedings of International Conference on Computational Thinking Education 2020* (pp. 53–56). The Education University of Hong Kong.

Li, Y., Schoenfeld, A. H., diSessa, A. A., Graesser, A. C., Benson, L. C., English, L. D., & Duschl, R. A. (2020). Computational thinking is more about thinking than computing. *Journal for STEM Education Research, 3*(1), 1–18. https://doi.org/10.1007/s41979-020-00030-2

Modeste, S. (2018). Relations between mathematics and computer science in the French secondary school: A developing curriculum. In Y. Shimizu & R. Vithal (Eds.), *ICMI Study 24, School Mathematics Curriculum Reforms: Challenges, Changes and Opportunities* (pp. 277–284). International Commission on Mathematical Instruction and University of Tsukuba.

Olofsson, A. D., Lindberg, J. O., Young Pedersen, A., Arstorp, A. T., Dalsgaard, C., Einum, E., & Willermark, S. (2021). Digital competence across boundaries - beyond a common Nordic model of the digitalisation of K-12 schools? *Education Inquiry*, *12*(4), 317–328. https://doi.org/10.1080/20004508.2021.1976454

Palts, T., & Pedaste, M. (2020). A model for developing computational thinking skills. *Informatics in Education*, *19*(1), 113–128. https://doi.org/10.15388/infedu.2020.06

Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. Basic Books.

Pérez, A. (2018). A framework for computational thinking dispositions in mathematics education. *Journal for Research in Mathematics Education, 49*(4), 424–461. https://doi.org/10.5951/jresematheduc.49.4.0424

Remillard, J. T. (2005). Examining key concepts in research on teachers' use of mathematics curricula. *Review of Educational Research, 75*(2), 211–246. https://doi.org/10.3102/00346543075002211

Shute, V. J., Sun, C., & Asbell-Clarke, J. (2017). Demystifying computational thinking. *Educational Research Review*, *22*, 142–158. https://doi.org/10.1016/j.edurev.2017.09.003

Swedish National Agency of Education (2018). *Curriculum for the compulsory school, preschool class and school-age educare 2011*. Elanders Sverige AB.

Tamborg, A. L., Elicer, R., Bråting, K., Geraniou, E., Jankvist, U.T., Misfeldt, M. (2023). The politics of computational thinking and programming in mathematics education: Comparing curricula and resources in England, Sweden, and Denmark. In B. Pepin, G. Gueudet, & J. Choppin (Eds.), *Handbook of Digital Resources in Mathematics Education*. Springer. https://doi.org/10.1007/978-3-030-95060-6_55-1

Wedman, L. (2020). *The concept concept in mathematics education: A concept analysis*. [Doctoral thesis, University of Gothenburg]. http://hdl.handle.net/2077/64096

Weintrop, D., Beheshti, E., Horn, M., Orton, K., Jona, K., Trouille, L., & Wilensky, U. (2016). Defining computational thinking for mathematics and science classrooms. *Journal of Science Education and Technology, 25*(1), 127–147. https://doi.org/10.1007/s10956-015-9581-5

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

Static, dynamic and interactive elements in digital teaching materials in mathematics: How do they foster interaction, exploration and persistence?

Anneli Dyrvold and Ida Bergvall

Department of Education, Uppsala University, Sweden

Contemporary comprehensive mathematics teaching material covering whole courses has developed substantially from the early versions that roughly were 'books as pdf' with some complementary online material. In teaching materials that are offered in online web portals (digital teaching platforms) a variety of dynamic and interactive elements can be utilised, offering new ways to engage with mathematics. Despite this recent development, the variety of affordances of the digital environment are utilised to a surprisingly small extent. The pros and cons with digital teaching materials in mathematics are debated, and publishers advertise with arguments about algorithms that lay out an ideal learning path and about joyful content. Critical for students' learning while working with teaching materials is however that they find it meaningful to use the materials, a persistence in the interaction with the materials, and furthermore that the willingness to explore mathematics remains. In this study students' interaction with digital teaching material with various kinds of dynamic and interactive elements supplementing the static parts in the presentation of new content is explored. Differences in students' attention to mathematical facts, essential in the problem solving, is captured using an eye-tracker. Analyses of differences in attentive behaviour depending on the kind of digital element that are used for presentation reveal that the type of digital element that students attend the least to is static elements. Differences in what is offered to and what is demanded from a reader when mathematical facts are presented using various digital elements is discussed and potential implications from the results are suggested.

Keywords: feedback, dynamic, attention, film, eye-tracking

ARTICLE DETAILS

LUMAT Special Issue
Vol 11 No 3 (2023), 103–131

Received 15 February 2023
Accepted 11 October 2023
Published 27 November 2023

Pages: 29
References: 55

Correspondence:
anneli.dyrvold@umu.se

https://doi.org/10.31129/LUMAT.11.3.1941

1 Introduction

The purpose of using teaching materials of different kinds is to support learning. From a student perspective however, this purpose is sometimes blurred by goals such as to get the right answer or to finish a section fast. Such goals can hinder learning, not least in a digital environment where active choices to 'open' or 'start' a particular part of the teaching material is needed to access all parts of the material. From this perspective teaching materials in print are beneficial because all content is displayed on the pages. A page-based layout, as in print, does however not surpass all affordances of the digital media. The digital media provides opportunities to visualise mathematics

 | CC BY 4.0

in new ways and to invite the reader to experience dynamic change (Dyrvold & Bergvall, 2023). Every learning opportunity does thus necessitate that students interact with the material, that they invest the required time, and make an effort to understand. Because of the particular features of digital teaching materials, knowledge about how students choose to use different dynamic features, and to navigate between different elements, is valuable to understand the potential these materials have to support learning.

The purpose of this study is to contribute to the understanding of what the presence of different dynamic elements in mathematics text means for students' interaction with the material. The current study focuses particularly on elements that are not static, that is *dynamic elements* that invite to interaction and/or present content that change over time (as e.g. films). The categorisation of different element types addressed in the research questions is presented in Table 1 in Method.

RQ1. How does students' interaction with mathematics differ depending on the dynamic element type used for presentation?

Not all parts of digital teaching materials are dynamic. In fact, a large share of the materials offered in digital teaching materials utilize static elements like text in print, but on screen. The second research question do therefore explore reading of static text in a context with included dynamic elements.

RQ2: Is there any difference in how students read static parts of an item due to presence of dynamic element types in other parts of the item, and if so, how?

1.1 Digital teaching material – in contrast to material in print

Teaching material in print has successively been complemented and partly also replaced with digital material. The use of digital teaching material is an intrinsic part of digitalization in schools (e.g., see European commission, 2021) and digital teaching material adds functions not possible to offer in print material. The potential of digital materials is particularly prominent in mathematics because of the possibilities to present relations between concepts and their representations dynamically. Besides the affordances of the digital media enabling interaction with and experience of mathematics in new ways, students' willingness to engage with dynamic material can also be beneficial in a learning situation. On the other hand, however, dynamic material may be distracting and can therefore potentially disrupt reading if displayed together with other material on screen.

The current study focuses on how students interact with digital materials that utilises different dynamic and interactive functions. Teaching materials, in print or digital, are supposed to offer rich learning opportunities to students and it is a reasonable argument that digital teaching material should in some manner offer better learning opportunities if the print material should be replaced. Previous studies who contrast print and digital teaching material highlight affordances of print materials such as the option to write in ink (Laughlin Davis et al., 2021), that text in print is easier to read (Abuloum et al., 2019), and the more natural sequencing of content in print (Gould, 2011).

The digital media, on the other hand, have many useful features that makes digital teaching material a very strong contestant to the traditional printed textbook. Such features, used in mathematics teaching material are for example, automated feedback (Van der Kleij et al., 2015; Stevenson, 2017), animations (Mamolo, 2019; McAlpin et al., 2019), dynamic visualisations of mathematical relations (Demir, 2018; Poon & Wong, 2017, Çeziktürk, 2020), drill-and-practice games (Beserra et al., 2019), and options to receive hints or access definitions (Stevenson, 2017; Arroyo et al., 2013). Analyses of contemporary digital teaching material do however reveal that much of the potential of the digital media is not taken advantage of (Glasnovic Gracin & Krišto, 2022; Mato-Vázquez et al., 2018; Dyrvold & Bergvall, 2023; Dyrvold, 2022).

1.2 Learning with dynamic and interactive material

Intervention studies contrasting learning with print material and digital material reveal different affordances of the digital media. For example, dynamic mathematics software (GeoGebra) has proven to be a powerful tool to help students think mathematically and to act like mathematicians and deduce hypotheses with the help of dynamic representations (Çeziktürk, 2020). Analyses of students' reasoning revealed that the connections that can be identified in multi representational mathematics software are crucial for the reasoning. Baccaglini-Frank (2021) also found that dynamic mathematics software contributes to students' development from a ritualistic discourse to a more explorative participation and thereby contribute to the construction of abstract mathematical objects. The usefulness of digital material is also revealed in a study evaluating evidence-based instruction, where results indicate that the instruction was helpful only for acquiring and maintaining mathematical knowledge if digital support was offered in the learning situation (Reinhold et al., 2020). Another affordance of digital materials is the possibility to offer virtual manipulatives, which for example can support understanding of patterns and concepts (Alagic, 2013). Furthermore, the option to digitally manipulate objects can increase students' motivation to persist. A meta-analysis of affordances of teaching materials offering dynamic representations of mathematical objects reveals that these materials can motivate students to persist at mathematical tasks. Other affordances revealed in the analysis are that these kinds of materials can encourage creativity, contribute to constraining students' attention to relevant content, and visualise relations between objects in relation to students' actions (Moyer-Packenham & Westenskow, 2013).

The digital media are useful for providing feedback of various kinds (e.g., Pinkernell et al., 2020; Ruthven, 2018) but rapid short feedback can also hinder learning if trial and error behaviour is used (e.g., Rezat et al., 2021b; Pinkernell et al., 2020; Nurmi & Jaakkola, 2006) whereas more elaborated feedback has proven useful for learning (Van der Kleij et al., 2015). The results of an in-depth analysis of individual learning with material offering automated feedback highlight that the design of the feedback is crucial to receive the desired effect (Rezat, 2021). Several studies reveal a positive relation between time spent reading digital teaching material and course grades (e.g., Junco & Clem, 2015), and furthermore, the time spent using certain parts of digital tools also predicted achievements in mathematics (Bokhove & Drijvers, 2012).

1.3 Interaction with dynamic and interactive teaching material

Digital material allows dynamic and interactive functions not possible to include in print, functions that also have proven to support learning. Students' interaction with these materials can however be shallow or without engagement and if so, the intended learning may be absent. But the opposite is also possible. It is argued based on the results of a meta-analysis that to foster motivation and persistence seem to be an under-emphasised affordance of virtual manipulatives (Moyer-Packenham & Westen–skow, 2013). An expected raised interest and willingness to engage is also a prevalent argument for use of digital games in learning. A study evaluating games in mathematics learning reveal that when a game mitigates anxiety, motivation and learning are enhanced (Huang et al., 2014) but there are also results revealing that a raised interest in a mathematical game-based activity can markedly decrease across sessions. Based on these results, Rodríguez-Aflecht et al. (2018) argues that game-based learning shall not be used to motivate students but rather used based on proven learning outcome.

If a chosen digital teaching material is engaging but does not scaffold learning, students' invested time is not fruitful, but the opposite is also true; if the teaching material has the best potential but students choose to not engage with it, the learning potential gets lost.

Design aspects of teaching materials are not new, but due to the diversity of dynamic functions and the modalities offered in digital media there is a huge increase in the kinds of learning opportunities that are possible to offer in a digital environment, in contrast to in print. This diversity gives rise to questions about which kind of digital features are useful in relation to which kind of mathematical activities and about the combination of dynamic and static elements in digital teaching platforms covering whole courses. Analyses of digital teaching platforms in mathematics reveal that dynamic and interactive functions are used only to a limited extent (Dyrvold, 2022; Mato-Vázquez et al., 2018), but also that when used such functions can extend the learning opportunities for example through activities where students can interact with the material or by adding a personal voice in a film (Dyrvold & Bergvall, 2023).

There is a growing research field focusing on digital teaching materials and many studies focus on some particular technology, investigating learning effects (e.g., Bray & Tangney, 2017;) or use in classrooms (e.g., Vahey et al., 2020). When it comes to teaching material that include a diversity of static and dynamic elements and students' use of such materials less is known, especially regarding students' interaction with the materials. Some recent studies focus on digital textbooks offering different kinds of elements (Pohl and Schacht, 2017; Brnic & Greefrath, 2022), others focus on students viewing time in different parts of online digital textbooks (O'Halloran et al., 2018; Kanwar & Mesa, 2022). The dynamic media used for such teaching material allows the reader to create their own reading path and to choose what to see and not. Reinhold et al. (2020) raises a question about ecological validity in studies about technology use in mathematics education, questioning whether results about potential benefits of technology use holds outside the experimental environment. For teaching material including a variety of static, dynamic, and interactive elements where students can choose whether to actively engage with all parts of the material, to what extent, and in what order, the use of digital features can be rather different than in an experiment. This dynamic and multimodal feature of digital teaching platforms in mathematics means that knowledge about their affordances, constraints and possible benefits must be built based on a variety of research studies. Knowledge built based on an accumulated bulk of research is nothing new, the issue here is that the variety of functions possible to include in digital teaching platforms means that separate studies can only comprise certain aspects of the material. For example, research reveals that rapid right/wrong feedback can foster trial-and-error behaviour whereas elaborated feedback and an option to choose to receive feedback step by step has been suggested as a better alternative (Rezat et al., 2021b; Heeren & Jeuring, 2019; Rezat et al., 2021a). If feedback options are included in a digital teaching platform with several other functions, however, the different types of feedback may be useful in relation to different parts of the material. For example, a few questions with rapid short feedback on the answer, in relation to theory presented as film, may foster active engagement.

Because of all options of reading paths and of what to visualise and interact with in digital teaching platforms, the opportunities to learn with the materials can differ substantially between users depending on their choices. Even materials with the best potential to support learning risk to miss their goal if the users do not invest time in the material and are not willing to explore and to be persistent. Designing digital teaching material that evidently aids all students in their learning is therefore a demanding task. Many studies do not reveal any convincing learning gains related to use of some digital teaching material despite thorough design of the materials. For example, materials developed with research based instructional design principles did not reveal a learning effect for all student groups (Reinhold et al., 2020). Another study synthesising results from 35 single-case studies on virtual manipulatives found that the use of commercially developed materials had a larger effect on students' mathematical accuracy after practice than researcher-developed materials (Shin et al., 2021). Together these results signal the complexity of both design of and learning with digital learning material.

In the current study our aim is to contribute knowledge about how students choose to interact with different kinds of dynamic and interactive elements. In digital teaching platforms students can choose what to display and the presence of one element is likely to have an impact on how other elements are read. Accordingly, the focus on interaction with different types of digital elements in the current study also includes analyses of students' distributed attention between various digital elements within a mathematics item.

2 Theory

2.1 Social semiotics

The current study is part of a research project with a social semiotic theoretical frame (see, Bergvall & Dyrvold, 2021). Taking this perspective on language means that various semiotic resources available in communication are considered as means for meaning making; these resources are signifiers whose meaning is dependent on the social and cultural context they are offered in (Halliday & Matthiessen, 2013; van Leeuwen, 2005).

The resources offered as means for meaning making do request different kinds of interaction from the reader. Halliday (1985) stresses that speech acts, also from text, are dialogic. Two commonly used speech acts are 'offering information' and 'demanding information'. These two speech acts are frequently used in mathematics teaching materials. For example, in theory sections, information is often offered as statements or agreed upon facts. In exercises, on the other hand, the speech act of demanding information is frequently used, in the form of questions or requests to engagement of some kind (see also Bergvall & Dyrvold, 2021). In a digital environment it is reasonable to also complement these categories with the speech acts related to 'offering goods and services' and of 'demanding goods and services' (cf. Halliday, 1985), because in a digital environment the dialog between the text and the reader includes other acts than offering or demanding information. For example, the reader may be demanded to use interactive elements to access information. All types of digital elements utilised in the section with facts (Figure 1, Method) in the items used in this study offer information. More dynamic and interactive digital elements do also, to a varying extent, demand 'digital acts', and in some cases also information in terms of responses, and the offered information is only provided after particular 'digital acts'. The theory constitutes an analytical framework in such a way that students' interaction with digital elements expressing different speech acts are compared.

2.2 Eye-movements and attention

Attention and gaze are strongly associated. The acuity in the centre of the fovea is far better than in the periphery, and accordingly, when some element is of particular interest it generally makes sense to fixate the gaze on it (e.g., see Pashler, 1999). Subjects can also attend away from where the gaze is fixated and shift the attention without moving the gaze. This divergence does not imply that shifts of attention and eye movements are unrelated, rather it is an expression of a person's ability to attend also to stimuli in the periphery. Two pairs of categories of attention and the relation between attention and gaze contribute to understanding the relation between the two. Firstly, visual attention can be either overt; that is a conscious act of physically directing the gaze to a stimulus, or covert; that is a mental shift of attention not related to a physical movement of the gaze. Typically, covert attention precedes spatial eye movements (e.g. Rai & Le Callet, 2018).

Secondly, attention can be either voluntary (endogenous attention) or involuntarily (exogenous attention) but despite different triggers, these types of attention often have similar perceptual consequences. Some exceptions indicate that endogenous attention is fairly flexible, in contrast to exogenous attention (Dugué et al., 2020). Endogenous attention emanates from the subject's mind whereas exogenous attention is caused by external stimuli. The current study is not designed to achieve data that discriminate between these categories, but these different kinds of attention are important to keep in mind when interpreting the results. For example, a student that interacts with a task can with an intention to grasp the offered content direct endogenous attention to a particular dynamic element. On the other hand, in interaction with dynamic elements particular visual displays may cause exogenous attention.

In summary, gaze fixations do not capture covert visual attention, because covert attention is the selective processing of information without change in gaze. Such covert visual attention does however precede a shift in visual attention to the particular location, and the coupling of attention and eye movements is mandatory. Interestingly, this relationship holds both for eye movements with exogenous control and with endogenous control (e.g., Hoffman, 1998). Attention and eye movements are not completely interdependent, but the relation is sufficiently prominent to be used as a foundation to learn about how texts are read. This widely accepted relation is often referred to as the eye-mind hypothesis.

3 Method

Data for this study was gathered using an eye-tracking analysis of 124 students in grade nine, drawn from four different schools in different parts of Sweden, while working on a set of five mathematics items. Data from 3 students were excluded due to poor calibration or missing eye-tracking data. One student rushed through all tasks in four minutes and data from this student was also excluded. This resulted in 120 students being included in the analysis. The group of participants do well represent the diversity in regular class in Sweden, with varying grades (for more information about the participants see also Dyrvold & Bergvall, 2023). The within-subjects design where all participants do all items and use all types of dynamic elements (Table 2) allows for inclusion of participants with different achievement levels without compromising the reliability of the study.

The students were informed that the overall purpose of the study concerned work with digital teaching platforms in mathematics and that the analysis would be carried out with eye tracking analysis to follow the participants' work with five mathematics items. The students were also informed that all participants are de-identified and that they can withdraw their consent at any time. All students who had reached the age of 15 gave consent to participate in the study (or their guardian in other cases).

3.1 Item design and included element types

The five mathematics items were designed to touch on areas of mathematics that are new to the students. The level of difficulty of the items was determined based on a review of Swedish textbooks in grade nine and thereafter adjusted in collaboration with two experienced teachers and textbook authors. The five items are about the inscribed angle theorem, maximum and minimum of quadratic functions, set theory, the relation between power and roots, and permutation and factorials. All five items had the same structure, consisting of four parts: introduction, task, facts essential for solving the task (hereafter called *Facts*) and answer options. Each of these four parts constitutes an area of interest (AOI). The eye-tracking analysis was built on how the student's gaze moved between these AOIs. These areas (that were not visualised on screen) were drawn closely around the constituents to minimise false positives (fixations that do not belong to the AOI) but not too small to avoid missed fixations that belong to the AOI. The shaded overlay around "Task" in Figure 1 visualises such an AOI.

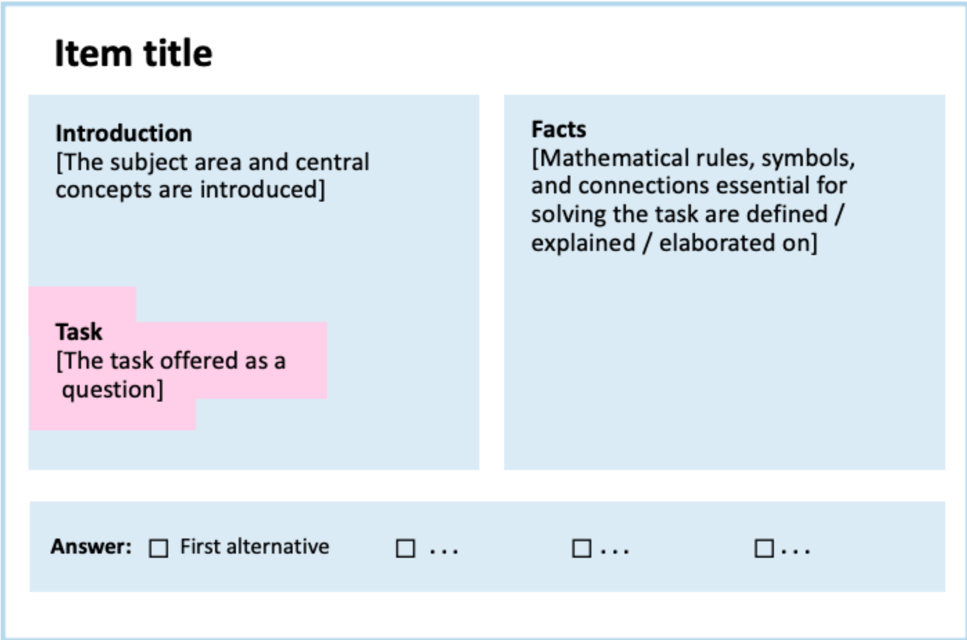


Figure 1. Basic design of the items and visualisation of an area of interest (AOI).

The presentation of the Facts using different element types (ETs) did not demand equal space, because different constituents were needed in the different versions. Because the number of fixations is likely to increase in larger AOIs (Holmqvist et al., 2011), these differences in occupied space were adjusted by somewhat enlarged constituents and row spacing. However, because different types of elements (Table 4) demand different kinds of constituents to present the Facts, equally sized AOIs would not have contained the same amount of information. Tightening constituents to achieve the same size if the AOI could instead diminish readability, and accordingly the best compromise between size and readability was sought in the design of the AOIs.

The Facts were designed in five versions, based on a typology of elements designed with an increasing interactivity and dynamism (Table 1) (Dyrvold, 2022). Elements are defined as a coherent part of a text that can contain both words, symbols, and images, where the components can be static and/or dynamic. Some of the elements are also interactive. Consequently, there are five versions of each item, and each version is dynamic and interactive to varying degrees depending on the element type (I–V) used in the Facts. All students were offered all five items in the same order, but the order of element type was varied (Table 2).

Table 1. Typology of elements used in Facts (Dyrvold, 2022).

Element type	Dynamic and interactive characteristics
I	static presence: elements are presented similar to a printed counterpart but on screen
II	opted presence: elements whose appearance is dependent on the readers' actions (e.g. click on a button)
III	dynamic presence: elements that change over time, typically content in a film or an animation
IV	dynamic feedback: elements that respond to and whose appearance is dependent on choices made by the reader (e.g. responses to students' answers)
V	continuous dynamic feedback: elements that change continuously over time depending on choices made by the reader (e.g. changing a slope in a coordinate system or moving geometric objects).

Because a central aspect of the method was to compare students' interaction with digital material depending on element types used for presentation, the different versions of Facts for a particular item should ideally offer the same information. Obviously, that is not possible because different semiotic resources contribute meaning in different ways, but as far as possible the offered meaning was kept similar. The typology (Table 1) allows for some variation within a particular ET. The same kind of dynamic and interactive features was however used within each ET in the study to increase consistency. For ET I, a static presentation of the Facts was used. For ET II the same static presentation was used but the Facts were hidden behind a button with the instruction "Click here to open". For ET III the same constituents (words, images, and symbols) were used as in ET I–ET II, but the content was presented in a film where the information appeared part by part simultaneously as a voice read the information out. Symbols were referred to by their names, but images were not referred to in the film. ET IV in all items consist of the same Facts as in the other ETs but the initial information was incomplete. In relation to some central aspect of the Facts two options of how to complete the facts were offered. The student was expected to 'tick' one of the alternatives and press "Check". If the response was "False" the correct alternative could be chosen to receive "Correct". In ET IV therefore, the students contributed to create correct Facts. In ET V a mathematical relation central in the Facts is visualised aided by a dynamic change of some kind. Typically, the students drag a slider that dynamically visualises for example values as 1, 2, 3. These dynamically changing values correspond to for example exponents in a mathematical expression and $2^2=4$; $2^3=8$ and so forth, is visualised corresponding to the values chosen by the slider.

The use of different ETs to present Facts means the reader is invited to different kinds of interaction depending on ET. Using ET I–ET III means the reader is 'offered information' and the demand of interaction is fairly small (i.e. to open a box [ET II] and to start a film [ET III]). In Facts where ET IV or ET V are used, parts of the information are instantly offered, but further information is offered only after some demands of interaction or information is met by the reader. For ET IV the reader must contribute to make the Facts complete, while receiving feedback regarding its accuracy, and for ET V the reader is demanded to move or drag constituents to be offered further information about the relation between the mathematical content in the moved spot and other central content (e.g. selecting a number of fruits [three] result in visualisation of the faculty [3!] for the number of ways the fruits can be ordered). A prominent difference between ET IV and ET V is that ET V offer information that dynamically change depending on the reader's action whereas ET IV changes instantly (not dynamically) in response to the reader's clicks.

To diminish potential carryover effects because of the within subjects' design, partial counterbalancing of the items was used. Five different items and five element types used to present the Facts gave 25 versions of the items (Table 2). These 25 versions were combined in ten different timelines where the order of the mathematics items was the same, but the element type used to present the Facts were altered.

Table 2. Items with different element types to present Facts combined in timelines.

Timeline	Item A	Item B	Item C	Item D	Item E
1:	Facts ET I	Facts ET II	Facts ET III	Facts ET IV	Facts ET V
2:	Facts ET II	Facts ET IV	Facts ET I	Facts ET V	Facts ET III
3:	Facts ET IV	Facts ET V	Facts ET II	Facts ET III	Facts ET I
…			…continued…		

Before working with the mathematics items, the students were provided individual information about the eye-tracking equipment, and an example item was shown on a screen. The students were informed that all items to be solved had the same type of layout and the main parts of the item (introduction, task, facts, and answers) were pointed out on the screen. The students were also informed about how the dynamic functions in the tasks work, either by being shown the dynamic functions by the experimenter or by having them try using the functions themselves. This instruction was provided in relation to a test item with a very easy task, an item that also was included before the five items during the eye-tracking. The inclusion of this test item before the five items contributed to diminish unwanted effects caused by distractions while the participants got acquainted with the digital environment.

3.2 Eye-tracking apparatus and measures

A portable eye-tracker was used because data collection was made at schools to make it equally accessible for all students to participate without a large effort. The eye-tracker uses binocular eye-tracking with a sampling frequency of 60 Hz with a precision of 0.10 degrees and accuracy of 0.3 degrees in optimal conditions. Tobii Pro Lab software (Tobii Pro AB, 2014) was used during data collection which provided good support to ensure the right distance between eyes and screen, by the provision of a virtual headbox to adjust to. The viewing distance was ≈ 68 cm and the monitor size was 15.6 inches.

Calibration procedure and AOIs. Students were allowed to wear glasses because the eye-tracker tolerated glasses well. If the calibration of a participant failed, it was repeated at maximum two times. Independent of the calibration results, all participants completed the assigned timeline but participants with bad accuracy values were excluded. The limit for exclusion was set at >1.0 degrees. This limit is reasonable because only four AOIs are used in every item and the distance between them is in most cases larger than 3 cm and never less than 2.5 cm (e.g. see Hessels et al., 2016).

Fixation filer. The threshold for fixations was set at 30 degrees/second. Adjacent fixations were merged with limits of a maximal angle between fixations of 0.5 degrees and maximum time between fixations of 0.5 degrees. Fixations shorter than 60 msec. were discarded.

3.3 Analysed variables and statistical analyses

The data from fixations used in analyses are accumulated fixation duration (AFD) within each of the AOIs and total number of fixations (NF) within each of the AOIs. Both AFD and NF are analysed because these variables represent similar but not the same kinds of attentive behaviour. Analysis of both variables is also valuable because similar results in relation to both variables is an indication of high reliability. Accordingly, the first two variables used in analyses are AFD and NF on the Facts. These variables are extracted in relation to the five types of element (ET) used to present the Facts. In addition to these variables, ratios of AFD on different AIOs in items with different types of elements (ET) used to present Facts are analysed. The ratio used in this variable is AFD on introduction (I) and task (T) in contrast to AFD on the Facts (F) ([AFD I+T]/[AFD F]). This means a larger ratio represents a minor share of AFD on the Facts in contrast to a smaller ratio. This variable is valuable in relation to potential high ADF on an ET because if the presence of a particular ET leads to increased attention and interaction, this increased attention is likely to be reflected also in a larger share of AFD on the ET in contrast to on the introduction and task (that is exactly alike in all versions of an item).

Because the focus in the current study is on various types of digital element types (Table 1), the analysed variables represent the five distinct different ET used to present the Facts. The design with five versions of every item and counterbalanced order of the items in timelines (Table 2) ensures that varying difficulty of the items depending on different content in Item A–E does not affect the reliability of the results.

To investigate whether there are differences in AFD, NF and ratios of AFD, between the five ETs a Kruskal-Wallis H test was used. A normality test revealed that the data do not meet the normality assumption which is why ANOVA was not used. Kruskal-Wallis H test is a non-parametric alternative to an ANOVA that can be used to determine if there are statistically significant differences between two or more groups of an independent variable on a continuous or ordinal dependent variable. In the current analysis, a null hypothesis suggesting no differences between the different ETs was investigated with a significance level of .050. A post hoc pairwise comparison was used to analyse differences between ETs (see Field, 2009). Effect sizes are calculated for pairwise comparisons, using Cohen's (1992) rule of thumb for effect sizes: r=0.10 small effect, r=0.30 moderate effect, and r=0.50 large effect.

4 Results

The first research question in this study concerns how students' interaction with mathematics differs with the element type used for presentation. Based on the eye-mind hypothesis, there is a connection between gaze and thought (Hoffman, 1998). By investigating how the accumulated fixation duration and total number of fixations on the different ETs differ, we gain information about students' interaction with the teaching content when it is presented using different ETs. A previous study has highlighted how different ETs have the potential to offer meaning in different ways (Dyrvold & Bergvall, 2023), which may have implications for the time and engagement students choose to spend on mathematics content when presented with different ETs.

Table 3. Descriptive statistics on accumulated fixation duration (AFD) and total number of fixations (NF) for Element type I – V.

Element type		AFD	NF
ET I	Mean	17092.46	50.15
	Median	14506.5	47
	SD	9794.373	26.65
ET II	Mean	17377.28	51.08
	Median	16288.5	47
	SD	10312.64	30.368
ET III	Mean	28424.48	80.29
	Median	25499	73.5
	SD	15775.58	41.136
ET IV	Mean	37419.12	106.6
	Median	33318.5	94.5
	SD	19807.28	61.098
ET V	Mean	31548.48	77.73
	Median	27639.5	71
	SD	19717.69	48.683

n = 120, Values for AFD presented are milliseconds

Table 3 shows that the mean and median values for both AFD and NF are lower in ET I and II compared to the other ETs. The Kruskal-Wallis H test showed that there was a statistically significant difference in AFD between the ETs, $H(4) = 147.6$, p= <0.001. There was also a statistically significant difference in NF between the ETs, $H(4) = 138.13$, p= <0.001, and thus that the null hypothesis should be rejected for both AFD and NF. The mean ranks for each ET used in the Kruskal-Wallis H test is displayed in Table 4.

Table 4. Mean rank of accumulated fixation duration (AFD) and total number of fixations (NF) on element type (ET).

	AFD		NF	
Element type	**Mean rank**	**Rank order**	**Mean rank**	**Rank order**
ET I	197.53	5	205.93	5
ET II	203.66	4	208.87	4
ET III	335.98	3	343.39	2
ET IV	415.90	1	422.85	1
ET V	349.44	2	321.47	3

The mean ranks for ET I and ET II are fairly similar and, a post hoc pairwise comparison (Table 5) reveals no differences between ET I and II, neither for AFD, nor for NF. These two ETs contain the same text, the only difference between the ETs is that the student must click to access the information in ET II, while it is statically presented in ET I. The largest effect sizes are found for the comparison between ET IV and ET I–ET II, and thus the contrast in number of fixations and fixation duration that stands out the most is on ET IV in compared to the more static ETs. All comparisons between ET I or ET II with the other ETs (ET III-V) reveal significant differences both in fixation duration and number of fixations. These differences signal that the more dynamic and interactive ETs (ET III-V) receive more attention both in form of accumulated fixation duration and number of fixations. The test also shows that there is no difference in fixations between ET III and ET V. One similarity between these ETs is that both, to a large extent, offer continuously dynamically changing content (see also Analysis and discussion).

Table 5. Pairwise comparison of accumulated fixation duration (AFD) total number of fixations (NF) on element types (ET).

	AFD					NF				
Comparison[a]	Test statistic	Z[b]	Sig.	Adj. sig.[c]	r[d]	Test statistic	Z[b]	Sig.	Adj. sig.[c]	r[d]
ET I- ET II	-6.137	-.274	.784	1.000	-.018	-2.942	-.131	.895	1.000	-.008
ET I- ET III	-138.450	-6.187	.000	.000	-.399	-137.458	-5.163	.000	.000	-.333
ET I- ET V	-151.917	-6.788	.000	.000	-.438	-115.538	-6.143	.000	.000	-.397
ET I- ET IV	-218.371	-9.758	.000	.000	-.630	-216.917	-9.693	.000	.000	-.626
ET II- ET III	-132.313	-5.912	.000	.000	-.382	-134.517	-6.011	.000	.000	-.325
ET II- ET V	-145.779	-6.514	.000	.000	-.420	-112.596	-5.032	.000	.000	-.388
ET II- ET IV	-212.233	-9.483	.000	.000	-.612	-213.975	-9.562	.000	.000	-.617
ET III - ET V	-13.467	-.602	.547	1.000	-.039	21.921	.980	.327	1.000	.063
ET III - ET IV	-79.921	-3.571	.000	.004	-.231	-79.458	-3.551	.000	.004	-.229
ET V- ET IV	-66.454	-2.969	.003	.030	-.192	-101.379	-4,530	.000	.000	-.292

[a]Each row tests the null hypothesis that the sample 1 and sample 2 distributions are the same.
[b]Standardised text statistic.
[c]The significance values was compared to significance values adjusted by the Bonferroni correction for multiple tests, with the same results.
[d] r is the effect size for pairwise comparisons (Field, 2018).

The second research question "Is there any difference in how students read static parts of an item due to presence of dynamic element types in other parts of the item, and if so, how?" was investigated using a ratio between accumulated fixation duration (AFD) on the static parts of the item, that is Introduction and Task, and the Facts. The assumption is that a low ratio means that the students spend more time on the Facts in relation to the rest of the task, which is interpreted as if the interaction with the Facts is informative enough to reduce the time spent on reading the introduction and solving the task. Table 6 shows descriptive statistics on the ratios, and the ratios for ET I and II are on average slightly higher than the ratios for ET III–V.

Table 6. Descriptive statistics on ratios of accumulated fixation duration (AFD) between areas of interest within items for different element types (ET).

Element type		AFD on Introduction and Task/ AFD on Fact
ET I	Mean	2.59
	Median	1.57
	SD	8.44
ET II	Mean	2.96
	Median	1.45
	SD	9.00
ET III	Mean	1.20
	Median	.85
	SD	2.39
ET IV	Mean	.77
	Median	.73
	SD	.43
ET V	Mean	1.49
	Median	.81
	SD	2.28

n = 120 (ETI. ET III. ET IV. ET V), n = 119 (ET II because AFD on Fact was 0 on ET II for one student)

The mean ranks of rations between accumulated fixation duration (AFD) on Introduction together with Task and Facts for each ET used in the Kruskal-Wallis H test is displayed in Table 7. The Kruskal-Wallis H test showed that there was a statistically significant difference in ratios between the ETs ($H(4) = 151.7$, $p < 0.001$). The ratio is between AFD on the three areas of interest: (Introduction + Task)/Facts.

Table 7. Mean rank of ratio for accumulated fixation duration (AFD) between areas of interest within items for different element types (ET).

Element type	Mean rank
ET I	411.51
ET II	392.77
ET III	235.77
ET IV	191.18
ET V	269.54

The result of a pairwise comparison of the mean ranks of ratios is presented in Table 8. A possible difference between ET I and the other ETs illustrates the difference between being offered only static text on screen and being offered also various types of dynamic elements in digital learning materials. The pairwise comparison indicates no difference between ET I and II, while ET III–V differ from the former in terms of ratios on ADF. All significant differences, except for between ET IV and ET V, have a moderate or high effect size. The ratios are lower for the more dynamic ETs, indicating that more time is spent on Facts compared to Introduction and Task. Table 3 shows that the students spend more time on the more dynamic ETs III–V and, moreover, Table 6 shows that the relative time of fixations on Facts in contrast to other text elements in an item is larger when ET III–V are used for presentation in contrast to ET I and II. A long AFD can, based on the eye-mind hypothesis, be interpreted as the student spending more time processing the mathematics content when ET is more dynamic.

Table 8. Pairwise comparison of ratio for accumulated fixation duration (AFD) between areas of interest within items for different element types (ET).

Comparison[a]	Test statistic	Z[b]	Sig.	Adj. sig.[c]	r[d]
ET IV - ET III	-44.583	-1.995	.046	.460	-.129
ET IV - ET V	-78.358	-3.507	.001	.005	-.226
ET IV - ET II	-201.590	-9.004	.000	.000	-.581
ET IV - ET I	-220.325	-9.861	.000	.000	-.637
ET III - ET V	-33.775	-1.512	.131	1.000	-.098
ET III - ET II	-157.006	-7.013	.001	.000	-.453
ET III - ET I	-175.742	-7.866	.001	.000	-.508
ET V - ET II	-123.231	-5.504	.001	.000	-.355
ET V - ET I	-141.967	-6.354	.001	.000	-.410
ET II - ET I	-18.736	-.837	.403	1.000	-.054

[a]Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.
[b]Standardised text statistic.
[c] Significance values have been adjusted by the Bonferroni correction for multiple tests.
[d] r is the effect size for pairwise comparisons (Field, 2018).

5 Analysis

All analyses in the current study reveal significant differences in fixations between the more static ETs, ET I–ET II, and the more dynamic/interactive ETs, ET III–ET V. There are however no significant differences in fixations between ET I and ET II. The reader needs to click-open the Facts in ET II. When opened, the Facts are identical for ET I and ET II. If there were a significantly higher number of fixations (NF) and/or accumulated fixation duration (AFD) for ET II that could be a positive effect of a perceived agency or 'investment' in the item, because actively opening the Facts is an investment of effort and engagement in the item. No such differences between ET I and ET II were identified and accordingly an option to limit the amount of information by hiding the Facts seems not to increase attention to the mathematical content.

In contrast to ET I–II, more attention (NF) is paid to ET III–ET V and there is also a larger share of fixations (AFD) on the Facts in items where ET III–ET V are used for presentation. At the least, these results reveal that the participants chose to attend a great deal to the Facts presented using ET III–ET V. A goal with teaching materials is to get the readers persistently engaged to learn with the material, and in this respect ET III–V meet the expectations.

An analysis of the speech acts that come into play when ET III–ET V is used gives another lens on the results. All three ETs *offer* information whereas ET IV–ET V *demand* more from the reader, that is, not only information but also, in Halliday's terms, 'goods and services' (in this context some digital acts). ET III differ from ET IV–V because ET III (film) demand a minor 'digital act' namely, to start the film, whereas ET IV–V demand both information and digital acts. Facts presented using ET IV as well as ET V do not offer sufficient information to solve the task without the readers' digital acts. ET IV offer incomplete statements and the reader needs to choose how to make the Fact complete, and thereafter check if the choice is correct via a button. If the wrong option is chosen, several trials is demanded to receive the correct information. This means there is not only a demand of digital acts from the reader, but also information because the reader has to choose options based on their knowledge. ET V utilizes drag-and-drop or sliders that when used results in visualization of dynamically changing content. Digital elements of type ET V are possible to use in a digital act without responding to demands of information by using the element type explorative. It is likely, however, that the reader also responds to demand of information, by choosing digital acts based on knowledge about the mathematics. An example of such an act of voluntary offering of information is when students pause a movement to visualise a particular mathematical relation or a critical value.

In summary, the analysis of speech acts reveal that ET IV demands the most from the reader. The analysis of fixation duration and number of fixations (AFD and NF, Table 3) also reveal the highest values for ET IV. The significantly higher values for ET IV indicate that the students choose to meet the demands of both digital acts and information, and that they thereby engage with mathematics. It is possible the students use trial and error behaviour (e.g. see Pinkernell et al., 2020) and choose not to use their own understanding while reading, but at the least ET IV demand information from the readers and they can choose to interact consciously.

6 Discussion

This study concerns five element types, and what their presence in mathematics text means for students' interaction with the material. The five element types are used to present Facts and are designed to reflect increasing dynamism and an increasing invitation to interact, from ET I to ET V (definitions of ET I–V can be found in Table 1, Method). Accordingly, the ETs differ in how information is offered and demanded (i.e. speech acts) from the student. Recall that ET I define static elements like print material but presented on screen. All other ETs have some dynamic features, not possible to offer in print.

Before a more in-depth discussion of the results, a reminder of what eye-tracking can and cannot say, is in place. Firstly, fixations do reliably represent attention, but the attention can be due to for example either complex constituents or attractive, joyful constituents. Secondly, many fixations and/or high accumulated fixation time represent persistent attention but does not reveal the extent to which the offered content is understood. The reasons why a participant choose to focus on the Facts can however vary. Persistent interaction can be caused by hard struggle to understand or by amusement from using a highly interactive ET. While interpreting these results it is therefore important to keep in mind that we study interaction, exploration, and persistence, not the reasons why the participants interact with the element types in particular ways. However, persistence is a prerequisite to grasp new and demanding content, and this study highlights more and longer fixations on content when some element types are used (ET III–ET V). The study do not analyse learning, but according to the theory behind eye-tracking (e.g., Pashler, 1999; Hoffman, 1998) the number of fixations and fixation duration is tightly related to either endogenous or exogenous attention. Based on the high interdependence between eye movements and attention the results can be used to understand more about how texts are read.

The results of the current study show that students spend time and attention on dynamic resources to a greater extent compared to when the information is offered as static text. Previous analyses of digital teaching materials in mathematics have shown that different dynamic elements have different potential to offer a reader meaning (Dyrvold & Bergvall, 2023). Regardless of the extent to which some dynamic resource can support meaning making, students must spend time reading and processing the information to take advantage of its potential. The importance of spending time has been highlighted in other studies revealing positive relations between achievement in mathematics and time spent on digital material as well as on using particular digital tools (Junco & Clem, 2015; Bokhove & Drijvers, 2012). Time is a blunt measure, but at the same time, it is easily understood that learning will not happen if the required effort is not invested. Accordingly, the indications of preference for dynamic and interactive elements revealed in the current study is worth reflecting over in the design of digital teaching materials.

ET III–V can be assumed to have an inherent demand of time, which has a direct bearing on the conclusions possible to draw based on the results. There are however reasons to assume that the extended attention on dynamic ETs is not solely caused by inherent demand of time in these ETs. The extent to which, for example ET V contributes to make mathematics accessible varies between the different items, but despite these differences, the differences in attention according to ET are significant. For example, in the easiest item about the inscribed angle theorem, the Facts of ET I is somewhat straightforward to interpret, whereas the version of the same item with Facts of ET V demand the reader to make connections between two dynamically changing angles. In more demanding items however, there are many meaning relations between constituents in the Facts, and in the more static versions of the Facts (ET I and ET II) the reader must understand all these relations, reasonably by circulating in the text, something that requires a great deal of attention. In ET V for such demanding items, on the other hand, the meaning relations are partly made apparent through dynamically changing objects in Facts. For example, in the item about permutations and factorials, the information in Facts can be rather hard to interpret for a student who is unfamiliar with the content. In static Facts, it is likely that the reader needs to read back and forth to fully understand the relations between quantity, permutations, factorials, and the meaning of the sign “!”. In Facts using ET V on the other hand, the reader is supposed to move fruit in and out of a casket and the permutations and the related factorial are instantly visualised. Accordingly, the consistent higher attention to Facts presented with ET V, possibly reflect an extensive exploration of the mathematics content rather than barely meeting the demand for interaction. These arguments support the claim that students attend more to content presented using interactive and/or dynamic elements (i.e. ET III–V) based on their own urge to gain understanding needed to answer the task, rather than based only on demands of time spent on the ETs because of their inherent features. In contrast to game-based learning that is criticised for only raising temporary interest, not to scaffold learning (e.g., Rodríguez-Aflecht et al., 2018), ETs that invites to exploration and thereby makes relations in mathematics more apparent have more of a scaffolding function. The scaffolding function is however conditioned by persistence.

One example of a dynamic element that students pay a great deal of attention to is ET III in the current study. A reason to interpret the time spent on ET III and as actively opted is that students in a pilot study (Dyrvold & Bergvall, 2022) expressed an experience that the film provided *more* information compared to the static ETs, despite the text on screen being the same. A conclusion is that the auditory resource in the form of a text reading voice in the film was experienced as "more" information that provided support for the reader. The pilot also revealed that participants chose to see the film repeatedly or to pause the film.

In items where ET IV is used, the student is requested to complete statements by choosing an option that makes a statement complete and correct. If the wrong option is chosen, the student needs to try again. Apparently, this kind of ET demands prolonged attention, especially if the wrong option is chosen in the first trial, and this demand of attention likely explains much of the fixations (NF and AFD) on Facts presented using ET IV. Because the results show that students pay a lot of attention to ET IV, one conclusion about the use of ET IV is that students are persistent enough in the strive to access the information needed to complete a task and that they are willing to make an effort and contribute to make offered Facts complete. From a didactical perspective, such investment of attention can potentially support learning more deeply compared to if the Facts were only read. While completing the facts through the options students need to evaluate and reflect over central concepts and mathematical relations. Being aware of wrong conclusions while learning new content is valuable (Boaler, 2015).

The downside of ET IV is that it can be used with a trial-and-error behaviour without learning gains. Previous research revel that automated feedback can be beneficial for learning but also that feedback in the form of short rapid response may not constitute the support needed, or that it even can be detrimental for learning if trial and error behaviour is used (Rezat et al., 2021a; Pinkernell et al., 2020). More elaborated feedback has been suggested to avoid such behaviour (e.g. Rezat, 2021; Heeren & Jeuring, 2019; Rezat et al., 2021b) but those suggestions are adapted to feedback in tasks. In the design of the items in the current study response options in ET IV, are used not as feedback in tasks but as guidance to complete facts needed to solve a task. This is quite a different way to use digital response options. A possibility is that the demand of 'digital acts' to complete the needed information is sufficient when it comes to this type of feedback as guidance to how the mathematical facts is to be understood.

For ET V on the other hand the demand of 'digital act' is minor, and the offered information is provided instantly. Because of the dynamic manner of ET V there are connections between constituents in the offered mathematics that the students need to grasp part by part. The demand of 'digital act' while reading content presented using ET V does therefore also entail a demand to be part in making sense of the offered information. A reader does of course always need to make sense of the offered meaning, but in contrast to static ETs where all information is present simultaneously, the meaning making while reading ET V can be considered a more constructive process. This process of meaning making can be an explanation to students' persistence while working with items with ET V as well as to their attention (high AFD and NF), because active learning engages. A previous meta-analysis about affordances of teaching materials with dynamic representations of mathematical objects highlight both that such materials can contribute to constrain students' attention to relevant content and that these materials can motivate students to persist at mathematical tasks (Moyer-Packenham & Westenskow, 2013). Together with other studies highlighting the benefits of dynamic material while learning mathematics (Alagic, 2013; Baccaglini-Frank, 2021) the meta-analysis strengthens the conclusion that ET V is likely to not only increase attention but also contribute to learning. The current study contributes to these previous studies by emphasising that dynamic elements have a great potential also in presentation of theory, not only in tasks which seems to be the most common way to use them (e.g. Dyrvold, 2022).

In conclusion, the results of this study highlight that the inclusion of more dynamic and interactive elements in digital teaching material can be useful to increase students' persistence and time spent on the mathematics content. The most prominent results from the statistical analyses are significantly higher attention (in the form of number of fixations and accumulated fixation duration) on Facts presented using dynamic and interactive element types (RQ1) compared to when static, or less dynamic, elements are used. Moreover, the increased attention on dynamic and interactive element types seem to reflect a persistence to sort out the information offered through these element types, either mathematical or linked to the dynamic function (the increase in accumulated fixation duration is allocated to the element types, not to the whole item) (RQ2).

In this way, our study elucidates that a more extensive inclusion of dynamic and interactive elements in digital teaching materials *can be* advantageous. This kind of resources are unique to digital teaching materials, and it is important to explore their function and advantages in teaching. The results can be useful as a guide when deciding in what manner these resources are beneficial, and when a static version reminiscent of a printed textbook is preferable. Thus, these results have implications for the development of digital teaching materials, but also for the teaching practice. When school leaders and teachers choose teaching resources, digital teaching materials offering dynamic and interactive elements must be considered based on an awareness that these elements can increase students' engagement in mathematics. Whether this engagement also leads to increased learning is not shown by this study and to investigate that further studies are needed.

Acknowledgements

This work was supported by the Swedish Research Council [grant number 2019-05005].

References

Abuloum, A., Farah, A., Kaskaloglu, E., & Yaakub, A. (2019). College students' usage of and preferences for print and electronic textbooks. *International Journal of emerging technologies in learning, 14*(7), 80–97. https://doi.org/10.3991/ijet.v14i07.9871

Alagic, M. (2013). Teacher candidates' perceptions of mathematical, cognitive and pedagogical fidelity of the "fill & pour" virtual manipulative. Proceedings of the 35th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education. University of Illinois at Chicago.

Arroyo, I., Burleson, W., Tai, M., Muldner, K., & Woolf, B. P. (2013). Gender differences in the use and benefit of advanced learning technologies for mathematics. *Journal of Educational Psychology, 105*(4), 957–969. https://doi.org/10.1037/a0032748

Baccaglini-Frank, A. (2021). To tell a story, you need a protagonist: How dynamic interactive mediators can fulfill this role and foster explorative participation to mathematical discourse. *Educational Studies in Mathematics, 106*(2), 291–312. https://doi.org/10.1007/s10649-020-10009-w

Bergvall, I., & Dyrvold, A. (2021). A model for analysing digital mathematics teaching material from a social semiotic perspective. *Designs for Learning, 13*(1), 1–7. DOI: https://doi.org/10.16993/dfl.167

Beserra, V., Nussbaum, M., & Oteo, M. (2019). On-task and off-task behavior in the classroom: A study on mathematics learning with educational video games. *Journal of Educational Computing Research, 56*(8), 1361–1383. https://doi.org/10.1177/0735633117744346

Boaler, J. (2015). *Mathematical Mindsets: Unleashing Students' Potential Through Creative Math, Inspiring Messages and Innovative Teaching*. Jossey-Bass.

Bojko, A. (2013). Eye tracking the user experience. A practical guide to research. Rosenfeld Media, LLC.

Bokhove, C. and Drijvers, P. (2012) Effects of a digital intervention on the development of algebraic expertise, *Computers & Education, 58*(1), 197–208. http://dx.doi.org/10.1016/i.comDedu.201 1.08.010

Bray, A., & Tangney, B. (2017). Technology usage in mathematics education research–A systematic review of recent trends. *Computers & Education, 114*, 255–273.

Brnic, M., & Greefrath, G. (2022). Distance learning versus face-to-face classroom learning: student achievement in conditional probability when using a digital textbook with integrated digital tools. *Twelfth Congress of the European Society for Research in Mathematics Education (CERME12),* Feb 2022, Bozen-Bolzano, Italy. hal-03748202

Çeziktürk, O. (2020). Understanding functional dependency on dynamic geometry systems (DGS): Napoleon and von aubel theorems on geogebra. *International technology and education journal, 4*(1), 15–21.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155–159.

Demir, M. (2018). Effects of virtual manipulatives with different approaches on students' knowledge of slope. *Journal of Interactive Learning Research, 29*(1), 25–50.

Dugué, L., Merriam, E.P., Heeger, D.J., Carrasco, M. (2020). Differential impact of endogenous and exogenous attention on activity in human visual cortex. *Scientific Reports, 10*(1), 21274–21274. https://doi.org/10.1038/s41598-020-78172-x

Dyrvold, A. (2022). Missed opportunities in digital teaching platforms: under-use of interactive and dynamic elements. *Journal of Computers in Mathematics and Science Teaching, 41*(2), 135–161.

Dyrvold, A., & Bergvall, I. (2022). The role of dynamic elements in digital teaching platforms: an investigation of students' reading behaviour. Proceedings of the Twelfth Congress of the European Society for Research in Mathematics Education. Presented at the 12th Congress of the European Society for Research in Mathematics Education, online, 2-6 February 2022. Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-474571

Dyrvold, A., & Bergvall, I. (2023). Digital teaching platforms the use of dynamic functions to express mathematical content. *Scandinavian Journal of Educational Research*. https://doi.org/10.1080/00313831.2023.2196555

Dyrvold., A., & Bergvall, I. (2023). Computer-based assessment in mathematics – issues about validity. *LUMAT: International Journal on Math, Science and Technology Education, 11*(3), 49–76. https://doi.org/10.31129/LUMAT.11.3.1877

European Commission (2021). *Blended learning for high quality and inclusive primary and secondary education: handbook,* Publications Office of the European Union. https://data.europa.eu/doi/10.2766/237842

Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5 ed.). SAGE.

Glasnovic Gracin, D., & Krišto, A. (2022). Differences in the requirements of digital and printed mathematics textbooks: Focus on geometry chapters. *CEPS Journal, 12*(2), 95–117. https://doi.org/10.26529/cepsj.1285

Gould, P. (2011). Electronic mathematics textbooks: Old wine in new skins? In *Proceedings of the 5th APEC-Tsukuba conference*, Japan.

Halliday, M.A.K. (1985). *An Introduction to Functional Grammar*. Arnold.

Halliday, M.A.K., & Matthiessen, C.M.I.M. (2013). *Halliday's Introduction to Functional Grammar (4th ed.)*. Routledge. https://doi.org/10.4324/9780203431269

Heeren, B., & Jeuring, J. (2019). Automated feedback for mathematical learning environments. In *ICTMT International Conference on Technology in Mathematics Teaching*. DOI: 10.17185/duepublico/48820

Hessels, R. S., Kemner, C., van den Boomen, C., & Hooge, I. T. C. (2016). The area-of-interest problem in eyetracking research: A noise-robust solution for face and sparse stimuli. *Behavior Research Methods, 48*(4), 1694–1712. https://doi.org/10.3758/s13428-015-0676-y

Hoffman, J. E. (1998). Visual attention and eye movements. In H. Pashler (Ed.), *Attention*, (pp. 119–154). University College London Press.

Huang, Y., Huang, S., & Wu, T. (2014). Embedding diagnostic mechanisms in a digital game for learning mathematics. *Educational Technology Research and Development, 62*(2), 187–207. https://doi.org/10.1007/s11423-013-9315-4

Junco, R., & Clem, C. (2015). Predicting course outcomes with digital textbook usage data. *The Internet and Higher Education, 27*, 54–63. https://doi.org/10.1016/j.iheduc.2015.06.001

Kanwar, P., & Mesa, V. (2022). Mapping student real-time viewing of dynamic textbooks to their utilization schemes of questioning devices. In *Twelfth Congress of the European Society for Research in Mathematics Education (CERME12).*

Laughlin Davis, L., Morrison, K., Zhou-Yile Schnieders, J., & Marsh, B. (2021). Developing Authentic Digital Math Assessments. *Journal of Applied Testing Technology, 22*(1), 1–11.

Mamolo, L. A. (2019). Development of digital interactive math comics (DIMaC) for senior high school students in general mathematics. *Cogent Education, 6*(1).

Mato-Vázquez, D., Castro-Rodríguez, M., Pereiro-González, C. (2018). Análisis de materiales didácticos digitales para guiar y/o apoyar el proceso de enseñanza-aprendizaje de las matemáticas. *@tic revista d'innovació educativa, (20).* 72–79.

McAlpin, E., Shilane, D., & Kalaycioglu, S. (2019). Dynamic versus static presentation formats, do they impact performance differently? *Journal of Computers in Mathematics and Science Teaching, 38*(1), 49–76.

Moyer-Packenham, P. S., & Westenskow, A. (2013). Effects of virtual manipulatives on student achievement and mathematics learning. *International Journal of Virtual and Personal Learning Environments, 4*(3), 35–50.

Nurmi, S., & Jaakkola, T. (2006). effectiveness of learning objects in various instructional settings. *Learning, Media & Technology, 31*(3), 233–247.

O'Halloran, K. L., Beezer, R. A., & Farmer, D. W. (2018). A new generation of mathematics textbook research and development. *Zdm, 50*(5), 863–879. https://doi.org/10.1007/s11858-018-0959-8

Pashler, H. E. (1999). *The Psychology of Attention*. MIT Press.

Pinkernell, G., Gulden, L., & Kalz, M. (2020). Automated feedback at task level: error analysis or worked out examples –which type is more effective? *Proceedings of the 14th International Conference on Technology in Mathematics Teaching – ICTMT 14.* https://doi.org/10.17185/duepublico/70767

Pohl, M., & Schacht, F. (2017). Digital mathematics textbooks: Analyzing structure of student uses. In G. Aldon & J. Trgalová (Eds.), *Proceedings of the 13th International Conference on Technology in Mathematics Teaching – ICTMT 13.*

Poon, K. K., & Wong, K. L. (2017). Pre-constructed dynamic geometry materials in the classroom--how do they facilitate the learning of "similar triangles"? *International Journal of Mathematical Education in Science and Technology, 48*(5), 735–755.

Rai, Y., & Callet, P. L. (2018). Chapter 3 - Visual attention, visual salience, and perceived interest in multimedia applications. In Academic Press Library in Signal Processing, Volume 6 (pp. 113–161). Elsevier Ltd. https://doi.org/10.1016/B978-0-12-811889-4.00003-8

Reinhold, F., Hoch, S., Werner, B., Richter-Gebert, J., & Reiss, K. (2020). Learning fractions with and without educational technology: What matters for high-achieving and low-achieving students?. *Learning and Instruction, 65*, 101264.

Rezat, S. (2021). How automated feedback from a digital mathematics textbook affects primary students' conceptual development: two case studies. *Zdm, 53*(6), 1433–1445. https://doi.org/10.1007/s11858-021-01263-0

Rezat, S., Fan, L., & Pepin, B. (2021a). Mathematics textbooks and curriculum resources as instruments for change. *Zdm, 53*(6), 1189–1206. https://doi.org/10.1007/s11858-021-01309-3

Rezat, S., Schacht, F., & Häsel-Weide, U. (2021b). Challenges of making sense of tasks and automated feedback in digital mathematics textbooks. In *Mathematics Education in the Digital Age* (pp. 168–184). Routledge.

Rodríguez-Aflecht, G., Jaakkola, T., Pongsakdi, N., Hannula-Sormunen, M., Brezovszky, B., & Lehtinen, E. (2018). The development of situational interest during a digital mathematics game. *Journal of Computer Assisted Learning, 34*(3), 259–268. https://doi.org/10.1111/jcal.12239

Ruthven, K. (2018). Instructional activity and student interaction with digital resources. In L. Fan, L. Trouche, C. Qi, S. Rezat, & J. Visnovska (Eds,). *Research on mathematics textbooks and teachers' resources: Advances and issues*. (pp. 261-276). Springer. https://doi.org/10.1007/978-3-319-73253-4_12

Shin, M., Park, J., Grimes, R., & Bryant, D. P. (2021). Effects of using virtual manipulatives for students with disabilities: Three-level multilevel modeling for single-case data. *Exceptional Children, 87*(4), 418–437.

Stevenson, C. E. (2017). Role of working memory and strategy-use in feedback effects on children's progression in analogy solving: An explanatory item response theory account. International *Journal of Artificial Intelligence in Education, 27*(3), 393–418. https://doi.org/10.1007/s40593-016-0129-5

Tobii Pro AB (2014). *Tobii Pro Lab* [Computer software]. Tobii Pro AB.

Vahey, P., Kim, H.-J., Jackiw, N., Sela, H., & Knudsen, J. (2020). From the static to the dynamic: teachers' varying use of digital technology to support conceptual learning in a curricular activity system. *Zdm, 52*(7), 1275–1290.

Van der Kleij, F., Feskens, R., & Eggen, T. (2015). Effects of feedback in a computer-based learning environment on students' learning outcomes: a meta-analysis. *Review of Educational Research, 85*(4), 475–511. Retrieved December 9, 2020, from http://www.jstor.org/stable/24753021

Van Leeuwen, T. (2005). *Introducing social semiotics*. Routledge.