Learning an Explanatory Model of Data-Driven Technologies can Lead to Empowered Behavior

A Mixed-Methods Study in K-12 Computing Education

Lukas Höper
lukas.hoeper@uni-paderborn.de
Paderborn University
Paderborn, Germany

Carsten Schulte
carsten.schulte@uni-paderborn.de
Paderborn University
Paderborn, Germany

Andreas Mühling
muehling@leibniz-ipn.de
Kiel University
Kiel, Germany
Leibniz Institute for Science and Mathematics Education
Kiel, Germany

ABSTRACT

Background and Context. One goal of K-12 computing education is to teach computational concepts that support learners in responsibly and competently using and evaluating digital technologies. However, recent research indicates that students struggle to make use of such concepts in everyday life. Additionally, research shows that people develop powerlessness and resignation about data-driven technologies, leading to passive user roles. This raises the question of how to support students' empowerment in navigating and shaping the digital world.

Objectives. This paper presents a study investigating how understanding concepts of data-driven technologies supports students' empowerment in everyday life. It involves developing an educational approach to support students in relating learned concepts to everyday experiences, called learning *explanatory models*.

Method. We have developed a Rasch-scaled instrument to measure understanding of data-driven technologies and motivation, intention, and empowered behavior in engaging with them in everyday life. Using this instrument, the study evaluates the explanatory model approach, which specifically supports such relations between concepts learned in computing and students' everyday experiences.

Findings. The results suggest that understanding of data-driven technologies according to our explanatory model leads to empowered behaviors in everyday interactions with such technologies. They also indicate improvements in students' understanding, intentions, and empowered behaviors in everyday life, while motivation did not significantly increase. We interpret that the approach supports students to make use of the concepts in everyday life and be more empowered in a digital world.

Implications. This paper demonstrates how the relationship between learning about data-driven technologies and the development of students' empowerment in everyday life can be examined. It shows how computing education can reduce students' resignations and powerlessness regarding data-driven technologies and support them in adopting more informed and empowered roles in navigating the digital world.

CCS CONCEPTS

• **Social and professional topics** → **Computing education**; **K-12 education**; *Student assessment*.

KEYWORDS

K-12, machine learning, data-driven technologies, empowerment, explanatory models, data awareness

ACM Reference Format:
Lukas Höper, Carsten Schulte, and Andreas Mühling. 2024. Learning an Explanatory Model of Data-Driven Technologies can Lead to Empowered Behavior: A Mixed-Methods Study in K-12 Computing Education. In *ACM Conference on International Computing Education Research V.1 (ICER '24 Vol. 1), August 13–15, 2024, Melbourne, VIC, Australia.* ACM, New York, NY, USA, 17 pages. https://doi.org/10.1145/3632620.3671118

1 INTRODUCTION

According to international K-12 computing education frameworks and research discussions, a goal of computing education at school level is to empower students to understand and navigate the digital world. For example, the ACM K-12 Computer Science Framework [49] or the Informatics Reference Framework for European computing education [16] emphasize enabling students to be informed citizens who understand digital technologies and the digital world. This involves understanding the role of CS in the world, participating in public discussion about CS topics, and competent and responsible use of digital technologies [49, p. 10] [16, pp. 4-5]. Given the prevalence of artificial intelligence (AI) and machine learning (ML) in everyday life, developing evidence-based teaching approaches about such data-driven technologies becomes an important topic in computing education. Understanding these technologies is necessary for competently using and evaluating individual and societal implications, reflecting on their role in the world, and shaping

ICER '24 Vol. 1, August 13–15, 2024, Melbourne, VIC, Australia

ACM ISBN 979-8-4007-0475-8/24/08
https://doi.org/10.1145/3632620.3671118

the digital world accordingly [48, 55, 79].[1] Therefore, teaching approaches should enable students to understand and reflect on the data-driven technologies from everyday life in order to become informed citizens empowered to engage in technological developments, instead of accepting such technologies as given.

In recent years, several approaches were developed to teach students about AI and ML concepts and enabling them to design and develop ML applications [38, 57, 68, 85]. As technologies using ML methods are driven by data practices, the role of data in these systems is essential [82, 84]. Consequently, teaching about data-driven technologies also involves teaching about data [38, 45, 57]. According to the goals of empowering to understand the digital world, students should be able to make use of these concepts learned in computing when encountering digital technologies in everyday life. They are required to relate and apply their knowledge of AI and ML to everyday situations. However, research indicates that students struggle to relate learned concepts about data and ML to everyday life and connect them to their everyday perspectives [e.g., 8, 10, 33, 88]. For example, studies report that students effectively learned basic ML concepts and could design ML applications but could not apply them to everyday situations [37, 88].

Students' motivations and intentions further reinforce this challenge of supporting students in using concepts learned in computing classes in daily situations, especially in the case of data-driven technologies. Research shows that many people develop feelings of resignation, powerlessness, and learned helplessness regarding the role of data in such situations, suggesting a clear passive role [e.g., 23, 24, 40, 44, 50]. However, empowering them involves enabling them to shift their focus to the inner workings of such technologies, rather than solely on their immediate user goal in these interactions (e.g., sharing personal experiences with friends). This research suggests that people may not see value in engaging with the inner workings of data-driven technologies or perceive themselves as competent enough to do so, which is at odds with the goals of empowerment.

This raises the need for approaches in K-12 computing education that support students in (1) relating computational concepts of data-driven technologies to their everyday experiences and (2) developing empowerment in a digital world accordingly. Therefore, this paper presents a study addressing these questions and focusing on bridging the gap between students' everyday user perspectives on data-driven technologies and an empowered, CS-informed perspective as a foundation for navigating the digital world.

Accordingly, we developed an educational approach termed learning *explanatory models*, in which students learn a model of data-driven technologies. This model is an educationally motivated composition of computational concepts of data-driven technologies, somewhat similar to notional machines designed to understand traditional programs in programming education [see 27]. This approach aims to enable students to apply the model in their everyday interactions with such technologies. In doing so, the model serves as a lens for reconstructing the inner workings of such technologies, which supports students in understanding and reflecting on these technologies and their individual and societal impacts. Thus, the core idea is to provide students with a model as a useful tool for analyzing, understanding, and reflecting on the data-driven technologies they encounter in everyday life.

To evaluate this approach, we developed an intervention for computing education. It allowed us to examine students' understanding of the explanatory model, their motivations and intentions to engage with the inner workings of everyday data-driven technologies, and their empowered behavior when interacting with such technologies. The study uses an instrument we have developed in our research project before. The evaluation of the intervention builds on data from $N = 93$ students from grades 9 and 10 (about 15 to 16 years old) in two European countries. The study examines the following research questions:

- RQ1. To what extent do students develop an understanding of data-driven technologies from their everyday lives?
- RQ2. To what extent does learning the explanatory model affect students' motivations and intentions to engage with the inner workings of data-driven technologies in everyday life?
- RQ3. To what extent does learning the explanatory model support students in empowered behavior in everyday interactions with data-driven technologies?

2 BACKGROUND

Regarding the aims of this study and the focus on the role of data in data-driven technologies, this section discusses prior work on teaching and learning about data-driven technologies, related research on powerlessness and resignation about their data practices, and underlying theoretical background.

2.1 Teaching and Learning about Data-Driven Technologies

For several years, much attention has been paid to how AI and ML could be taught at the school level, emphasizing AI literacy and AI education. Some meta-reviews already provide overviews in this fast-growing research field [e.g., 15, 38, 57, 75]. The literature shows several ideas for effectively teaching AI and ML concepts [e.g., 25, 42, 55, 88, 94]. This involves, for example, developing ML models or applications [25, 88, 94]. One promising trend is using and developing educational tools that offer easy options for designing ML applications [e.g., 42, 94]. These tools are designed to make learning ML accessible even without prior programming experience. They allow students and teachers to focus on the conceptual aspects, such as training and testing ML models, rather than programming at code level, following the principle of 'low floors, high ceiling, and wide walls' [74], making learning about AI and ML a realistic and achievable goal for young students. In addition, a few approaches also delve into critical perspectives, for instance, discussing ethical and societal implications of AI systems to enable students to become informed and responsible consumers of AI technologies [55]. Another example is ethically redesigning AI technologies from everyday life, encouraging students to envision a future world with AI they desire [22]. While distinguishing different levels of learning outcomes about AI (e.g., technical aspects of ML models and ethical considerations), Rizvi et al. [75] analyzed literature on K-12 AI education. They consider four levels: engines,

[1] In this paper, "competent use" refers to interactions with digital artifacts that are based on an understanding of its inner workings and technical aspects. Similarly, we use the term of "engaging with digital technologies" to refer to engaging with how they work and what they are for.

models, applications, and social and ethical considerations [see 75]. Their literature review showed that while many approaches focus on ML models and respective AI applications, they often neglect societal and ethical considerations.

As data-driven technologies rely on the data used to train the models, the role of data is essential in such technologies and strongly influences their behavior [see 82, 84]. Hence, AI education also includes relations to data-related concepts and data literacies [38, 45, 57]. Such data literacies involve a wide range of approaches, which primarily focus on practical data handling skills, such as collecting, analyzing, and evaluating data (for a comprehensive overview of different perspectives, see [33]). Additionally, there are calls for broader perspectives like critical reflections on data practices and their societal and ethical implications [e.g., 8, 71].

Regarding personal involvement in learning data or data-driven concepts, research demonstrated positive effects when contextualizing in personal, relevant situations or perspectives. For example, Register and Ko [73] found that learning with personal data can support understanding ML concepts. Similarly, Bilstrup et al. [8] highlighted the importance of linking data and ML concepts to personal data, facilitating students' engagement and prompting reflection on data practices in their personal lives.

In summary, teaching about data-driven technologies in K-12 encompasses learning basic data and AI concepts, understanding AI technologies, working with data, developing AI applications, and critically evaluating individual and societal implications. Despite the importance of understanding and evaluating digital technologies from a user perspective according to the aims discussed at the beginning, a designer-oriented viewpoint is omnipresent, neglecting perspectives like reconstructing these technologies and their impact [65, 75].

2.2 Everyday User Perspectives and Designer Perspectives

Students frequently use data-driven technologies in everyday lives, such as social media applications, streaming services, and search engines. Through these interactions, students form perceptions, gain experiences, and may develop misconceptions about data-driven technologies, as examined in several studies [62, 63, 67, 76]. Building on these perceptions and experiences within constructivist learning approaches can effectively scaffold students' understanding of computational concepts and their relation to everyday life. However, the predominant focus on designer-oriented views in K-12 AI education creates a discrepancy in students' roles in everyday interactions with such technologies, as they usually have no designer roles.

Various versions of a continuum between user and designer roles are described in the literature [21, 28, 77, 79, 80]. This continuum delineates the difference between using digital technologies and actively designing and shaping them. For instance, regarding the role of digital technologies in everyday life, Rushkoff [77] posits the dichotomy of "program or being programmed" to argue that when not being empowered to create and design digital technologies, individuals must adapt themselves to technologies made by others. From this perspective, K-12 computing education aims at facilitating transitions from passive user roles to active designer roles, that is, empowering them to shape the digital world [21, 49, 79].

Drawing on examining students' attitudes towards computing, Schulte and Knobelsdorf [80] advocate for integrating professional use perspectives when learning to design digital technologies to support different pathways in computing. When confined to designing digital artifacts, students face challenges translating abstract concepts learned in computing classes into everyday life contexts, as evidenced by research findings. Studies indicate students' limited understanding of the role of data and data practices in everyday digital technologies [e.g., 10, 13, 36, 45, 71, 72, 84]. When learning about data practices or data-driven technologies, studies report that students struggle to relate this to their everyday lives [10, 33, 37, 88]. For instance, students may find it challenging to recognize their role as data sources when using data-driven applications [33] and to grasp such concepts of data on a concrete personal level [10]. Additionally, a study revealed that while students learned about ML concepts and designing ML applications, they did not critically reflect on the data-driven practices in their everyday lives [88].

Taken together, relating computational concepts to everyday situations and using them when encountering data-driven technologies poses a significant challenge for students.

2.3 Motivation and Intention as Drivers for Using Concepts in Everyday Life

When examining whether students effectively engage with the inner workings of data-driven technologies in everyday life and apply the computational concepts learned in computing classes, motivational and intentional factors play a crucial role in addition to understanding the concepts. Drawing from theories of human behavior, the Rubicon model is frequently referenced, stating that actions (e.g., using the concepts in everyday life) are determined in four phases of goal setting and goal striving: predecisional, preactional, actional, and postactional [1, pp. 486-490]. According to this model, motivation and intention are crucial for actions, as crossing a boundary is required to initiate behavior after goal setting. While motivation pertains to the choice of goals (i.e., the decision-making), intention refers to committing to realizing this goal and making a respective plan [1, 41]. Therefore, this study incorporates motivation and intention to engage with the inner workings of data-driven technologies in everyday life, with the theoretical foundation described below.

We adopt the expectancy-value model in this study, as widely used in educational research for understanding motivation [92]. It elucidates the factors influencing why individuals choose, initiate, and direct their engagement in activities, such as engaging with the inner workings of data-driven technologies. Motivation, according to this model, comprises two components. Firstly, expectancy concerns beliefs and confidence in one's ability to complete or succeed in a specific task or activity [92]. This aligns with the concept of self-efficacy proposed by Bandura [3] [see also 92], which refers to individuals' beliefs in their capacity to act successfully in a given situation and achieve desired outcomes. Self-efficacy is specific to particular contexts or actions [3, p. 42]. Research indicates that it significantly influences peoples' decisions and behavior, thus predicting effort, engagement, and achievement in the targeted activity or behavior [3, 81]. Secondly, as described by Wigfield and Eccles [92, 93], the value component comprises intrinsic value, attainment value (or importance), utility value, and cost (see for further distinctions: [93]). Intrinsic, attainment and utility values positively influence motivation, whereas cost diminishes individuals' willingness to engage in the activity [93]. These value components are related to individuals' interests in contents as the perception of value is one of the dimensions in interest development [see 64]. Perceiving such a value in learning can influence students' long-term interest and engagement in computing topics [64].

In addition to motivation, we consider intention, which refers to the commitment to and planning for goal achievement. One type of behavioral intention is implementation intention, which involves planning to realize a goal in specific situations, that is, committing to a particular course of actions and behavior [1, 35]. These intentions often take the form of if-then plans, specifying when, where, and how to realize the goal [35]. Several studies provide evidence that implementation intentions positively influence goal attainment and behavior initiation [for an overview, see 35], thus leading to the formation of behaviors and habits [90]. Additionally, research indicates that motivation can predict individuals' intentions and choices of activities [91], although intentions do not necessarily imply motivation for the given activity [2].

2.4 Further Challenges Due to Perceived Powerlessness and Resignation

Students' experiences with data-driven technologies profoundly influence their motivations and intentions. Negative experiences with data-driven technologies from daily life can hinder students from effectively applying computational concepts in everyday situations, particularly when they perceive this as useless. In the realm of data-driven technologies, especially with an emphasis on the role of data, students' perspectives, feelings, and attitudes regarding data-driven practices were investigated.

The theory of learned helplessness, a well-established concept, is particularly relevant in contexts of data-driven technologies. It concerns perceived independence of one's actions and the respective outcome, fostering a belief of futility of one's behavior [58]. This sense of helplessness extends across affective and cognitive dimensions. In this respect, several conceptual frameworks emerged in privacy research: privacy apathy [40, 50], privacy cynicism [44], privacy fatigue [19], privacy helplessness [18], and resignation [24]. They denote feelings of powerlessness or resignation concerning personal data and individuals' control over data privacy and data practices involving those practices in data-driven technologies.

This can lead to what is known as the privacy paradox, describing a dissonance between individuals' privacy concerns and actual privacy behaviors [5, 6]. Recent research indicates that students are concerned about data practices but often feel powerless to make a difference, lack understanding of the implications, and want to learn more about it [23, 78]. Studies report that learning about data practices does not necessarily alter students' intentions to disclose personal data [8] or alleviate feelings of powerlessness [40].

Overall, this research indicates that students' feelings of powerlessness and resignation toward the role of data in digital technologies contradict the goals of empowerment discussed at the beginning. Such resignation and powerlessness would increase the risk of being a passive consumer, potentially preventing students from engaging with these technologies' inner workings, even if they had experiences in designer roles. Therefore, computing education needs approaches to support students in using the concepts in everyday interactions with data-driven technologies and encourage them to actively explore the inner workings of such technologies rather than passively using them. To target this gap, we developed an approach that addresses students from their user perspective and equips them with an explanatory model to support understanding and evaluating the inner workings and influences of everyday data-driven technologies.

3 EDUCATIONAL APPROACH OF EXPLANATORY MODELS

This section outlines the educational approach of learning *explanatory models* as the foundation for this study. Its core idea is to explicitly teach students a model designed from an educational perspective - in contrast to a model taken directly from CS. It should provide an explanation for computational concepts and digital technologies. Thus, the approach aims to support students in learning concepts in a way that helps them form mental models useful to understanding data-driven technologies they encounter in everyday life. It is similar to teaching notional machines in programming education used to explain the behavior of programs during their execution (see for an overview: [27]). However, while notional machines are taught as a vehicle until students understand the real behavior of programs, explanatory models are taught as an end goal themselves.

We have designed an explanatory model focused on the role of data in data-driven technologies, providing a specific lens on these technologies (a prior version was published in [45]). We hypothesize that learning this model supports students in using it as a lens or tool to analyze and evaluate data-driven technologies they interact with in everyday life, that is, to understand their inner workings and reflect on the interactions with these technologies. Below, we describe this explanatory model and then the pedagogical idea of how it was taught in this study.

3.1 Explanatory Model for the Role of Data in Data-Driven Technologies

As described earlier, studies demonstrated that contextualizing learning about data and data-driven technologies within personal perspectives increased personal engagement [8, 73]. Thus, we designed the model from students' everyday user perspective. Notably, the explanatory model makes a specific framing in terms of an educational reconstruction of computational concepts and does not primarily aim to draw on specific data or AI concepts from the discipline as they are often rather abstract [see 26]. In this regard, the model reconstructs the role of data in data-driven technologies, offering a lens to understand the role of data in these technologies and to be integrated into students' everyday lives. The model describes both technical aspects of data collection and processing and underlying purposes of these data practices. It is illustrated in Figure 1 and briefly described below.

How is data collected in data-driven technologies? During interactions with data-driven technologies, data about users and the interactions hold particular significance as they form the basis for many data-driven features. This data is collected through various methods, as outlined in several distinctions [56, 70, 71]. For instance, data can be categorized as follows: provided data that the user actively creates; observed data gathered through observation and recording, of which the user is not necessarily aware; derived data generated by directly processing existing data; inferred data generated by probability-based processing [70]. Thus, users intentionally provide data, which we call *explicitly collected data*. In contrast, other data is collected through observation, tracking, and data processing alongside the interaction, which we call *implicitly collected data*.

How is the data processed? Data-driven technologies utilize practices from data science and ML to process collected data, generating data models such as ML models or user models, which are created by user modeling [for overview to user modeling, see 12]. This paper focuses specifically on *data models about users*, known in the literature as user models, digital footprints, digital shadows, data doubles, or digital doppelgänger [e.g., 9, 51, 86]. These models, derived from explicitly and implicitly collected data, are usually continually refined during interactions with data-driven technologies (e.g., based on user modeling and profiling techniques). Additionally, they may include sensible information, even if not provided by users themselves [36, 66]. For example, predictive analytics methods can predict sensitive personal data based on specific predictive models [66]. The data models can then be used to predict preferences and future behavior [51, 95]. This is, for instance, especially in contexts like recommender systems, which use these predictions for filtering methods. While the explanatory model is designed from users' perspectives, the data models about users are crucial due to their individual implications, such as privacy concerns and targeting. However, ML models are also interesting to consider due to potential impacts on interactions, leading to bias and fairness issues.

For which purposes is the data processed? Data collection and processing serve various purposes, often driven by commercial interests. From a technical perspective, data processing aims to provide different features and generate outputs during interactions that users may recognize. These are called *primary purposes* in our model. However, providers often have additional intentions beyond the immediate output generation. For instance, data is processed to predict users' behaviors for feature adaptation [e.g., 66, 86, 95]. These predictions may be leveraged for targeted advertising or, at another level, to influence users' behaviors and emotions [52, 83]. We call them *secondary purposes*, which are often not apparent to users, sometimes intentionally obscured or overshadowed by primary purposes [e.g., 14, 95]. Distinguishing these purposes allows students to evaluate data practices of data-driven technologies in a nuanced way, as it often depends on the different purposes.

3.2 Teaching and Learning the Explanatory Model

Taken again the user-designer continuum discussed earlier (see Section 2.2), the overall goal of this approach is to support students in navigating along this continuum from a user to a designer role of actively shaping digital technologies and develop future technologies [21, 77, 80, 84]. Consequently, it aims at making abstract computational concepts about data-driven technologies accessible and applicable to students' everyday lives, enabling them to understand and reflect on such technologies (i.e., becoming informed as first steps on the continuum). This facilitates competently and responsibly navigating the digital world, ultimately empowering them to contribute to shaping the digital world [48, 55, 65, 79].

Thus, the approach teaches students an explanatory model and enables them to use it as a lens for data-driven technologies. This study's intervention for teaching this explanatory model is oriented to context-based learning and the idea of semantic waves [60]. Context-based learning is well-established in science education [7, 34], but also used in computing education, such as implemented in problem-based learning [39, 69]. Accordingly, we use an authentic and meaningful real-life context to support students in relating and applying the concepts to everyday life [34]. Therefore, learning the explanatory model is embedded in an everyday context of interacting with a data-driven technology, allowing students to reconstruct the role of data in such situations. To counteract the potential limitation of acquiring only context-specific knowledge [see discussion in: 39], the abstract concepts of the explanatory model (e.g., data models about users) are subsequently introduced and applied to the context through cyclical processes of decontextualizing and recontextualizing. This aligns with the idea of semantic waves, which describes transitions between levels of semantic density (i.e., condensation or complexity of knowledge) and semantic gravity (i.e., relation to a context) [60]. For example, this concerns transitioning from explaining an abstract concepts, moving on to a concrete real-world example, and then returning to the abstract concepts (see for more details: [60, 89]). In doing so, students' everyday experiences are addressed to support students in relating the explanatory model to their everyday lives.

4 METHOD

This section outlines the procedure, participant details, intervention, instrument for data collection, and subsequent data analysis.

4.1 Procedure

The ethics review board of the corresponding author's institution granted approval for this study, which is part of a broader design-based research project on developing the data awareness framework [see 45, 46]. This study examines how learning the explanatory model supports students in relating it to their everyday lives and fosters their empowerment in navigating the digital world.

Given that students encounter data-driven technologies in their everyday lives from an early age, our research project focuses on secondary school students in grades 6 to 10 (12 to 16 years old). We have developed two interventions for computing education in secondary schools, which mostly differ in the chosen context: one for grades 6 and 7 and another for grades 8 to 10. These interventions

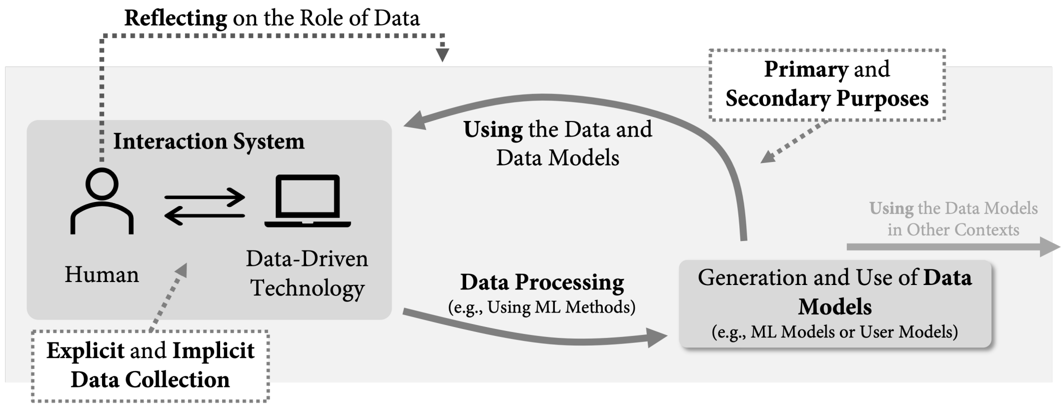

Figure 1: The explanatory model that provides one possible lens on the role of data in data-driven technologies [see also 45].

serve as a basis for evaluating the explanatory model approach. Our evaluation of the approach is conducted through a pre-post-test design, with a questionnaire as the measurement instrument (see details in Section 4.5). This instrument was developed as part of the broader research project, including multiple rounds of piloting and evaluation on a broader sample of students. Its development is subject to another article that we are preparing, but we briefly summarized the results of the instrument evaluation in Section 4.5.2. The study presented here focuses on the intervention for grades 8 to 10 (outlined in Section 4.3) and involves 93 students. Figure 2 provides an overview of this design.

The sample selection was based on recruiting teachers from collaborations in prior projects and teacher training courses given by one of the authors. One of the authors introduced the teachers to the study, questionnaire, and intervention. After obtaining written informed consent from the students and their legal guardians, teachers introduced the questionnaire during their computing classes, where students completed it using a digital survey tool. For students unfamiliar with terms like data, teachers were permitted to explain it briefly to aid comprehension of the questionnaire items. The teachers conducted the intervention in their classes, with clarification on the intervention and materials provided in meetings before and during the study. After the intervention, students completed the questionnaire again. The intervention and surveys took place during regular computing classes and lasted 6 to 8 lessons of 45 minutes each, typically over 3-4 weeks.

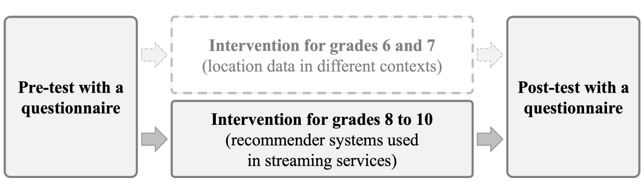

Figure 2: Overview of the study design. This paper evaluates the explanatory model approach according to the intervention for grades 8 to 10.

4.2 Participants

The study presented here involved 93 students (67 males, 24 females, 2 non-binary) in grades 9 and 10 across Germany and Switzerland. These students came from 6 classes at 4 different secondary schools in rural and urban areas and from different types of schools (differing in general student achievement levels). Table 1 provides a respective overview of the classes. Many of these students have already received computing education in school, as computing was introduced as a compulsory subject in the official curriculum of the participating schools a few years ago. Consequently, many students probably had some prior experience with data and AI concepts.

Table 1: Sample description

Class	Participants	Country	School	Grade
A	10	Germany	A	10
B	15	Germany	A	10
C	11	Germany	B	10
D	18	Germany	C	10
E	20	Germany	C	9
F	19	Switzerland	D	9

4.3 Intervention: Role of Data in Services with Recommender Systems

The intervention adopts the explanatory model approach. It uses an exemplary data-driven digital artifact from students' everyday lives to support the relation to everyday experiences. Our choice was movie streaming services using recommender systems. Details of the intervention are described below and summarized in Figure 3.

The intervention consists of four phases. In the *first phase*, the example of a movie streaming service is used as a starting point, raising the question of how such a system generates personalized recommendations. This introduces the core idea of recommender systems. Working in small groups, students develop initial ideas for generating such recommendations and discuss which personal

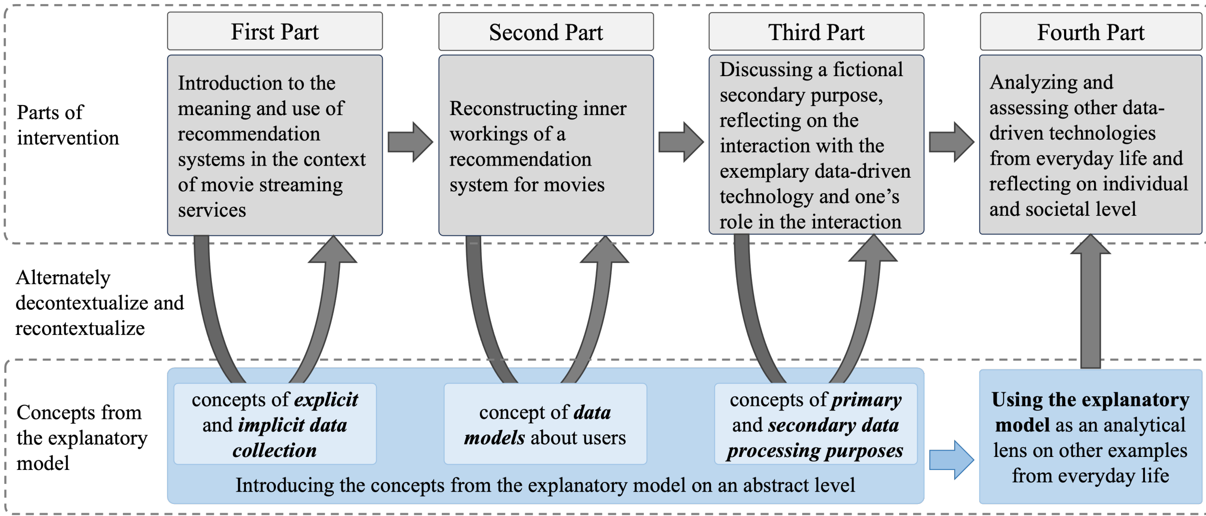

Figure 3: Overview of the intervention used in this study, implementing the approach described in Section 3.2.

data could inform deciding which movies to recommend. Building on this activity, the idea of data collection is decontextualized, and the concepts of *explicit* and *implicit data collection* are introduced. Students then recontextualize these concepts by mapping their ideas to these types of data collection.

The *second part* delves deeper into the black box of movie recommender systems. Students are given a prepared Jupyter notebook, offering an opportunity to look under the hood. This Jupyter notebook contains a Python module that we have developed, which generates personalized movie recommendations for students based on rating data from real people. Through a step-by-step exploration, students uncover details of data collection (e.g., collecting movie ratings and tracking user behavior) and the subsequent data processing for generating personalized movie recommendations. This includes an introduction to k-nearest-neighbor as an example of an ML method, enlightening students on how ML models could be developed based on user data. By decontextualizing the idea of collaborative filtering, the concept of *data models about users* is introduced. Students then recontextualize this concept within the given context and explain data models about users and their role in predicting user interests based on similar user ratings.

The *third part* deals with secondary data processing purposes of recommender systems. It begins decontextualized with an introduction to the concepts of *primary* and *secondary purposes*. Students recontextualize these concepts by using them to summarize the previous part (focus on primary purposes) and brainstorm potential secondary purposes. Through a panel discussion, students engage in dialogue regarding the hypothetical secondary purposes of a personalized paywall based on predictions of users' movie preferences. From diverse viewpoints, students deliberate on this idea and reflect on the responsible use of such technologies. They also reflect on their own role within these interaction systems and explore various individual and societal implications of streaming services (e.g., regarding filter bubbles). These evaluations consider the different types of purposes.

During the *fourth part*, students engage with the role of data in data-driven technologies from their everyday lives that use recommender systems. Working in groups, they select such an artifact and explore it according to the concepts of the explanatory model. This process allows students to apply the concepts to other everyday examples, supporting experiences in reconstructing the role of data in data-driven technologies and reflecting on the interactions. For instance, this part may involve examining specific apps the students use. After analyzing and evaluating the role of data in these examples, students discuss possible advantages and disadvantages. Thus, critical reflections are encouraged.

4.4 Hypothesized Model for this Study

As previously outlined, this study examines students' understanding, motivation, intention, and empowered behavior. As these four components are interrelated, we use a structural equation modeling approach (detailed in Sections 4.5 and 4.6). Accordingly, the study builds on a hypothesized model encompassing four components and their assumed relations, forming the basis for the questionnaire (see Section 4.5). This theoretical model is shown in Figure 4 and elucidated below.

The four components are defined as following:

- **Understanding** of the role of data in data-driven technologies according to the explanatory model,
- **Motivation** to engage with the role of data in data-driven technologies in everyday life,
- (Implementation) **Intention** to engage with the role of data in data-driven technologies in everyday life, and
- **Empowered Behavior** in terms of being empowered to make sense of inner workings of data-driven technologies from everyday life, instead of being a passive consumer.

Between these four components, we assume several relations. Based on the discussions about motivation and intention (see Section 2.3), the motivation component indicates whether students have chosen the goal of engaging with the role of data in everyday

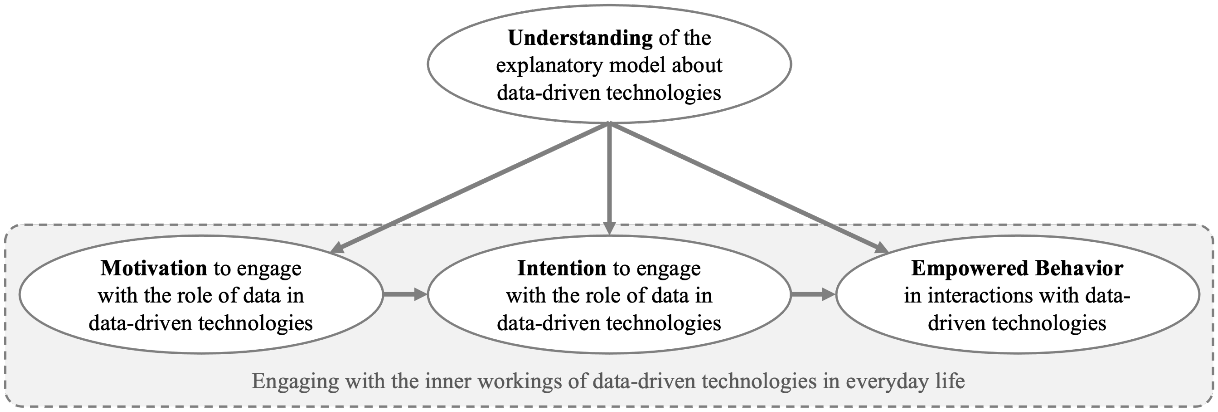

Figure 4: The hypothesized model with four components as covered in the questionnaire.

data-driven technologies. Additionally, implementation intention indicates whether students plan to realize this goal of engaging with the role of data. Given the research on motivation and intention discussed earlier, motivation can lead to implementation intention (see Section 2.3). Taken together, motivation and intention indicate the extent to which students want to engage with the inner workings of data-driven technologies (i.e., to use the concepts in everyday life). Subsequently, the intention may translate to engaging with the inner workings of data-driven technologies in everyday life (see Rubicon model in Section 2.3), that is, being empowered to become more informed about these technologies. Hence, we assume a predictive chain of relations from motivation to intention to empowered behavior. Furthermore, as discussed regarding resignation and powerlessness (see Section 2.4), understanding could affect the motivation and intention to engage with the inner workings of data-driven technologies. Additionally, it is hypothesized that this understanding supports students in becoming informed and competent users in interactions with such technologies, thereby influencing empowered behavior.

4.5 Instrument for the Data Collection

Based on the theoretical model, we developed a questionnaire reflecting measurement models for the respective components. The questionnaire was used as both pre- and post-test. It consists of four parts covering the components of the hypothesized model (see Figure 4). Additionally, it asks for an ID and demographic information (grade, school, gender). Students completed the questionnaire using a digital tool, which begins with personal information and then presents the four parts and their items in a randomized order to mitigate a question order bias. In this section, we briefly describe the questionnaire and the results of its evaluation (Table 2 shows the final items)[2].

4.5.1 Description of the Questionnaire.

Context in the questionnaire. The questionnaire captures a situation that students are familiar with from everyday life to examine whether they could relate the concepts to everyday life. We chose "using apps" as the context due to the students' widespread use of apps and the prevalence of data-driven functionalities (e.g., in social media apps). By contextualizing the questionnaire around interactions with apps, it eliminates the need for more concrete perceptions of AI and mitigates effects of misconceptions about AI [e.g., 62]. The understanding part refers to the situation of installing a new social media app, including a description of a fictional app with a news feed (i.e., it uses a recommender system). The motivational, intentional, and behavioral parts are more broadly framed around using apps and engaging with the role of data in them.

Part 1. Understanding the explanatory model. The explanatory model described before proposes a way to characterize the role of data in data-driven technologies, whose aspects are meant to be computational concepts. This study involves examining students' understanding according to this explanatory model. This part of the questionnaire, therefore, includes items about three themes: (a) data collection, (b) generation and use of data models, and (c) data processing purposes. As far as we know, there is no measurement for this perspective on data-driven technologies. Thus, we developed items accordingly. Initially, a subset of the authors developed a set of items drawing from the explanatory model, prior research, and literature on data-driven technologies. The initial draft included open-ended and true-false rating items. An iterative process of reviews and discussions with different computing education researchers led to several modifications, increasing items' content validity. Additionally, we have revised some items during the piloting process to improve their comprehensibility (see piloting process in Section 4.5.2). The items used here included two open-ended items regarding understanding of data collection and data processing purposes. These stem from a previous study, where they had shown to be useful and usable from grade 6 on [see 45]. Additionally, it included 18 items regarding data models that students should rate as true or false.

Part 2. Motivation and intention to engage with the role of data. The motivation and intention items were mostly adapted from existing validated scales. The motivation items are rated on a 7-point Likert scale ranging from "strongly disagree" (1) to "strongly agree" (7). We incorporated self-efficacy, intrinsic value, importance, and

[2]The full questionnaire is published as external supplementary material: https://doi.org/10.5281/zenodo.11609812

cost based on the expectancy-value theory of motivation. *Intrinsic value* is similar to the construct of intrinsic motivation because both refer to engaging in an activity driven by interest and enjoyment [92]. Therefore, we used a validated scale for intrinsic motivation consisting of seven items to assess the intrinsic value [17] (see description and validation in: [59, 61]). The items were adapted according to the previously described activity, similar to what has been done for other studies [see 17]. The perceived *importance* was also covered using a validated scale with six items [17] (see also: [59, 61]). The items were adapted to the activity addressed in the questionnaire, taking into account other exemplary adoptions [see 17]. To measure a perceived *cost* as part of the value component, we adapted a scale developed by Flake et al. [30]. As we focused on activity, we chose the task effort cost subscale and adapted this to the activity mentioned in our questionnaire. This subscale has five items. As far as we know, there has not been an appropriate *self-efficacy* scale related to whether people believe they can understand the concepts of our explanatory model for data-driven technologies. Therefore, following Bandura's recommendations for developing self-efficacy scales [4], we developed six items that cover different aspects of the explanatory model. Similarly, due to the specific activity, we designed our own items for the *implementation intention* subscale. Such items can be "if-then" statements rated according to one's commitment [2, 35]. These items can indicate that students have these behavioral intentions. In literature, using two or three items related to specific activities is suggested to predict intentions [29]. Hence, we developed two items rated on a 7-point Likert scale ranging from "I will certainly not do that" (1) to "I will certainly do that" (7).

Part 3. Behavior in interactions with data-driven technologies. Given the difficulty of observing and assessing students' everyday behavior, we included this part to gain insights into students' perspectives on their interaction behavior and whether these indicate a sense of empowerment. As scant research operationalizes "empowerment" for quantitative measurements, we decided to include an open-ended item to explore whether students' answers reflect informed and empowered behavior or rather powerlessness. Recognizing the potential for social desirability bias when asking about their behavior (see privacy paradox: [5, 6]), we contextualized the open-ended item as offering advice to a friend on engaging with the role of data in apps. The answers could reveal their views on such an interaction behavior.

4.5.2 Overview of the Prior Instrument Evaluation. The previously described instrument was developed in our broader research project. The prior piloting process and evaluation of the instrument are summarized below.

Piloting. We conducted multiple rounds of piloting with teachers and students to improve the questionnaire's comprehensibility for young students. This process involved individual thinking aloud with a 6th-grade student, thinking aloud with a teacher, and testing the design with a whole 6th-grade class. In particular, we observed some difficulties in item formulations, resulting in some improvements and removing one item from the motivation scale. Due to this process and the further analysis of the instrument (see description below), we observed that many young students struggled with negations of items, resulting in avoidance of negated items in the final questionnaire.

Prior instrument evaluation. In our research project, we collected data from 398 secondary school students in grades 6 to 10 across two European countries (210 males, 177 females, 11 non-binary), including the 93 students who participated in the study presented here. These 398 students are from 25 classes from 17 different schools. Based on these data, we evaluated the instrument, following a mixed-methods approach, including content analysis, Rasch modeling, and a typical structural equation modeling approach [32]. The results of this analysis is briefly summarized below.

Firstly, the qualitative data were analyzed following a thematic qualitative content analysis [53, pp. 69-88]. We developed separate code systems for each of the three open-ended items presented in Table 3. Data sessions and inter-rater tests on randomly chosen material were used to ensure the coding was valid and reliable. The agreement ranged from substantial to almost perfect between the different code systems.

Secondly, the rating items of the understanding part were evaluated with a Rasch modeling approach, as these were assessment items that we believed to capture a unidimensional construct. Based on the item format, we used a dichotomous Rasch model; model fit was examined using Andersen's LR-test on total item-set level and the Wald-test on single item level. After excluding some items based on the Wald-test results, the LR-Test provided for the final scale a p-value of 0.50, indicating a good fit. The outfit values of the items are on average 1.02, and the infit values 0.98, indicating that the resulting scale is usable for assessment.

Thirdly, we used confirmatory factor analysis to test the measurement models of motivation and intention with a five-factor model. After excluding two items with small factor loadings, the final five-factor model showed a good fit according to fit criteria suggested by Hu and Bentler [47] ($\chi^2(179) = 272.49, p < 0.000, CFI = 0.971, TLI = 0.965, RMSEA = 0.042, SRMR = 0.047$). The standardized factor loadings range from 0.62 to 0.91. The implementation intention subscale has acceptable reliability ($\alpha = 0.79$), while the other subscales have (very) good reliabilities (α between 0.83 and 0.91).

Fourthly, we tested the structural equation model (SEM), which showed a good model fit ($\chi^2(515) = 681.589, p < 0.000, CFI = 0.963, TLI = 0.959, RMSEA = 0.031, SRMR = 0.054$) [see fit criteria: 47]. The SEM is shown in Figure 5.

Using expert interviews and the SEM analysis, the prior evaluation of the instrument suggested adequate face and construct validity. Overall, the prior evaluation showed that the instrument is a sufficient measurement for the study presented in this paper.

4.6 Data Analysis for this Study

According to the research questions, this study's data analysis focused on evaluating the effects on students' understanding, motivations, intentions, and empowered behavior.

The analysis of the qualitative data from the questionnaire follows a thematic qualitative content analysis [53, pp. 69-88]. For each of the three open-ended items, we developed separate code systems. The codes for the data collection and processing items (see part 1 of the questionnaire) were deductively defined, while those for the

Table 2: The final set of the questionnaire items (translated from German into English)

Subscale	Items
Understanding data collection	SocialNet would collect and use data about you. Please note what data SocialNet is likely to collect about you.
Understanding data processing purposes	Please note for what purposes SocialNet might use the data about you.
Understanding data models	SocialNet creates digital doppelgänger of its users. Rate the following statements about digital doppelgänger as true or false. • When I use SocialNet, my digital doppelgänger becomes more and more precise. • A digital doppelgänger describes a person in its entirely. • The provider of SocialNet can use my digital doppelgänger to predict what I am likely to do on SocialNet in the future. • The provider of SocialNet can also use digital doppelgängers beyond SocialNet. • A digital doppelgänger contains only true information about the user. • In a social network like SocialNet, my digital doppelgänger consists of my profile page and my posts. • My digital doppelgänger influences what SocialNet displays to other users. • Social networks can influence user behavior. • Letting a social network collect data about myself is my own decision, which does not concern other people. • The way I use a social network can affect other people. (*Note.* This includes both true and false statements.)
Intrinsic value	Decide how much you agree with these statements: When I am engaging with the data collection and processing of an app ... I am enjoying it very much. / I find it is fun to do. / I would describe it as very interesting. / I think it is quite enjoyable.
Importance	Decide how much you agree with these statements: When I am engaging with the data collection and processing of an app ... I put a lot of effort into it. / I try very hard to understand it. / I try very hard to do it. / it is important to me to understand it.
Cost	Decide how much you agree with these statements: Engaging with the data collection and processing of an app ... demands too much of my time. / requires me to spent too much energy. / takes up too much time. / is too much work. / requires too much effort.
Self-efficacy	Decide how much you agree with these statements: When I am engaging with the data collection and processing of an app, I would be sure to find out ... what data I create in the app. / what data the app collects about me incidentally while I am using it. / what data is required to use the app's features. / for which other purposes the app processes data about me. / what choices I have to influence the data collection and processing. / how the app might influence me, for example, by selecting the posts I see.
Implementation intention	Decide how much you agree with the following statements: • When a new app is released soon, I will think about what data this app would collect and process about me before installing it. • The next time I use an app on my smartphone, I will look at what data the app collects about me and what it processes it for.
Empowered behavior	Imagine a friend of you gets a new smartphone and is going to install apps. Which advice would you give about whether and how to engage with the role of data in such apps?

behavior item (see part 3 of the questionnaire) were inductively generated. Coding manuals for the deductive code categories were derived from the explanatory model (i.e., covering the types of data collection and the kinds of data processing purposes defined in Section 3.1) and refined during data sessions. The inductive code categories for the behavior item were generated by exploring the data and discussing them in data sessions. This analysis resulted in four main code categories described below in more detail (see Section 5.3). Table 3 presents the final code systems and respective examples.

One of the authors coded all qualitative data, while a second coder worked on about 20% of the data randomly chosen for each of the three items. The inter-coder reliability was calculated using the Brennan and Prediger Kappa [11]. The strength of agreement was almost perfect for the data collection item ($\kappa = 0.83$), almost perfect for the data processing item ($\kappa = 0.81$), and substantial

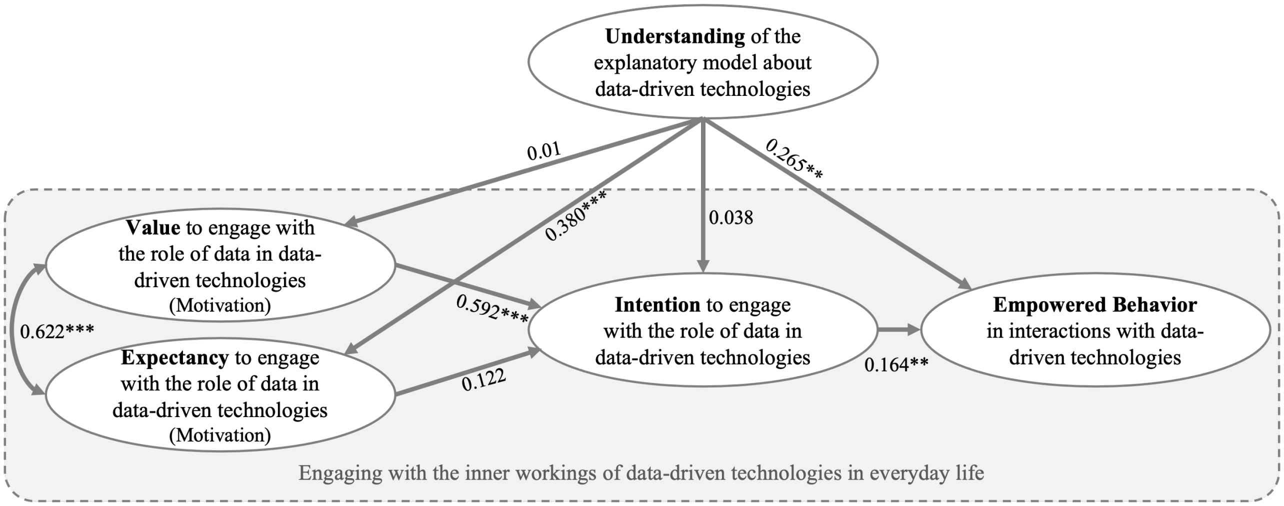

Figure 5: Results from the SEM analysis according to the hypothesized model. (Significant paths are marked with ** for $\alpha < 0.01$, and * for $\alpha < 0.001$.)**

for the behavior item ($\kappa = 0.72$) [54]. The coded data were also quantified and further analyzed using quantitative methods.

As the items and scale values are ordinal scaled but not all normally distributed according to the Shapiro-Wilk test, we used the Wilcoxon signed-rank test to examine pre-post differences (two-sided with $\alpha = 0.05$). This analysis examined changes in students' understanding, motivations, intentions, and empowered behavior. The effect size r of results were interpreted as small between 0.1 and 0.3, medium between 0.3 and 0.5, and large above 0.5 [20].

5 RESULTS

The results are presented according to the research questions.

5.1 Students' Understanding of the Role of Data

The data from the open-ended items about the data collection and the data processing purposes were coded deductively according to the aspects of the explanatory model. Table 3 provides an overview of the codes for these items and the respective frequencies. The data indicate that prior to the intervention, only some students already had a nuanced perception of the data collection. However, most of the students (80.7%) had no understanding of the data collection or could only describe the more obvious explicitly collected data (e.g., "my name and email address"). After the intervention, nearly the half of the students could describe examples of implicitly collected data (e.g., "it collects how often I am online and what I click there") or mentioned both types (43%). According to the prior conceptualization of these types of data collection (see Section 3.1), the four codes are hierarchically ordered and thus ordinal scaled. The pre-post difference is significant with large effect size according to the Wilcoxon signed-rank test ($W = 154.5, p = 0.00005, r = 0.614$). This indicates that the students developed a more nuanced understanding of the data collection of data-driven technologies.

Regarding the data processing purposes, the secondary purposes are mostly mentioned rather superficially (e.g., "personalized advertising"), while the primary purposes require a deeper understanding of the inner workings of data-driven technologies. As an example for a primary purpose, a student mentioned "deriving data about interests that can be used to filter the posts in my news feed." In the pre-test, most of the students had no ideas about the purposes (22.6%) or only mentioned superficial ideas about the secondary purposes (41.9%), while about a third of them could describe primary purposes or both types (31.2%), indicating a deeper understanding about technical aspects of the data practices. After the intervention, many students could describe primary purposes or both types (65.6%). Since this data are also ordinal scaled, we used the Wilcoxon signed-rank test. This provides that the pre-post differences are significant with medium effect size ($W = 508, p = 0.0013, r = 0.412$). This indicates that the students developed a better understanding of the purposes for which the data are processed in data-driven technologies. In particular, the results suggest that many students developed a better understanding of the inner data practices.

Based on the Rasch-scale part for understanding data models in data-driven technologies, the scale values can be calculated as sums so that the range of possible scores is up to 10. In the pre-test, students' scores are on average 6.09 ($sd = 2.29, md = 6$), while they are 6.98 ($sd = 2.78, md = 7$) in the post-test. According to the Wilcoxon signed-rank test, this pre-post difference is significant with medium effect size ($W = 890, p = 0.0018, r = 0.356$). This indicates that many students developed a better understanding of the data models about users, how they are generated, and which role they play in interactions with data-driven technologies.

Taken together, the results for the three understanding parts suggest that the students developed a more nuanced understanding of the role of data in data-driven technologies. As the exemplary contexts from the intervention and the questionnaire are different, these results indicate that the students developed an understanding of the concepts from the explanatory model.

Table 3: Frequencies of codes in the qualitative data from the pre-post data ($N = 93$)

Code	Examples from students	Number of students in pre-test	Number of students in post-test
Data Collection:			
(0) no answers	-	13 (14.0%)	3 (3.2%)
(1) explicit data collection	my name; my text messages to friends	62 (66.7%)	50 (53.8%)
(2) implicit data collection	online status; posts I have viewed	4 (4.3%)	5 (5.4%)
(3) both types of data collection	(examples for both types)	14 (15.1%)	35 (37.6%)
Data Processing:			
(0) no answers	-	21 (22.6%)	18 (19.4%)
(1) secondary purposes	personalized advertising	39 (41.9%)	14 (15.1%)
(2) primary purposes	display posts in my news feed; check the age restriction	18 (19.4%)	32 (34.4%)
(3) both kinds of purposes	(examples for both kinds)	15 (16.1%)	29 (31.2%)
Empowered Behavior:			
(0) no ideas or resignation	you can't do anything	23 (24.7%)	20 (21.5%)
(1) ideas when to shift the focus	if you have to register for an app with a profile; if you are asked for personal information about yourself	30 (32.3%)	16 (17.2%)
(2) ideas on what to focus	look at what data the app wants to track from you; check whether the app is trustworthy	23 (24.7%)	31 (33.3%)
(3) ideas what could be done	hide your email address (there are providers for this); try to give as little information about yourself as possible; block trackers	17 (18.3%)	26 (28.0%)

5.2 Students' Motivation and Intention towards Engaging with the Role of Data

The students were surveyed about their motivations and intentions to engage with the role of data in everyday data-driven technologies. Figure 6 provides an overview of the scale statistics for the motivational and intentional factors.

Regarding the intention to engage with the role of data in everyday apps, students rated the items with an average of 2.82 ($sd = 1.39, md = 3$) in the pre-test and with 3.34 ($sd = 1.8, md = 3.5$) in the post-test. The pre-post difference is significant with medium effect size according to the Wilcoxon signed-rank test ($W = 799.5, p = 0.0023, r = 0.356$). This indicates that the intervention encourages many students to commit to the plan of engaging with the role of data in everyday interactions with data-driven technologies. Thus, more students rated such an engagement as more relevant after participating in the intervention. The scale values of the motivational factors are a bit better in the post-test, but the differences are not significant. Students' motivation for this engagement is moderate to low, especially regarding the value factors of motivation.

Moreover, according to the SEM (see Figure 5), the self-efficacy component of motivation is affected by understanding the concepts ($\beta = 0.380, p < 0.000$), while the effect of understanding on the value component is not significant ($\beta = 0.010, p = 0.899$). Regarding the relation between motivation and intention, only the value component significantly affects students' intentions ($\beta = 0.592, p < 0.000$).

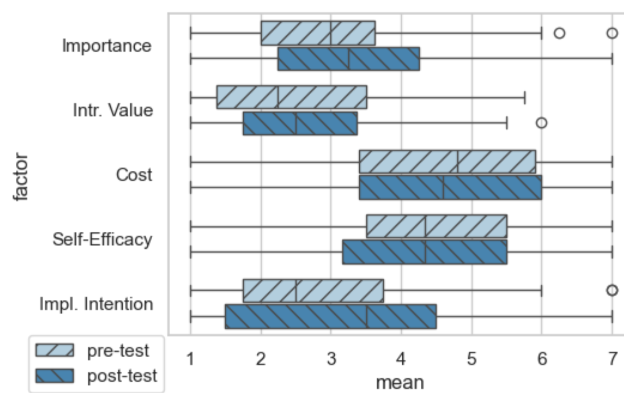

Figure 6: Boxplots for the motivational and intentional factors grouped by pre- and post-test

5.3 Students' Behavior of Engaging with the Role of Data

The students were asked when and how to engage with the role of data in data-driven technologies in everyday life. The inductive qualitative analysis of this item focused on identifying different levels of being informed and empowered. This resulted in four code categories (see Table 3).

The first code is that students had no ideas about the question or stated to be resigned and having a state of surrender. The second code was assigned when students mentioned having ideas for reasons when to shift one's focus away from the immediate goal of why one is interacting with a digital artifact, that is, situations in which they value engaging with the role of data (e.g., "if you are asked about personal information"). The third code was given when students described ideas of with what to engage (i.e., where to focus on) to get information about the inner workings and data practices (e.g., "look at what data the app wants to track from you"). The last code refers to answers with ideas for possible actions in such interactions with data-driven technologies (e.g., "block trackers").

Regarding further analysis, students were assigned to one of these code categories as a score. They got the highest code contained in the answer (i.e., the answer can also contain segments with lower codes but no higher ones). Thus, scores for empowered behavior are given, which are ordinal scaled. Table 3 reports the order of the codes, code examples, and frequencies.

The scores are on average 1.37 ($sd = 1.05, md = 1$) in the pre-test, and 1.68 ($sd = 1.1, md = 2$) after the intervention. The pre-post difference is significant with moderate effect size according to the Wilcoxon signed-rank test ($W = 488, p = 0.025, r = 0.306$). These differences are visualized in Figure 7, showing the transitions of students from pre-test to post-test according to the scores. It can be seen that many students made improvements towards the post-test, even if a few students were assigned a lower score in the post-test than in the pre-test. This indicates that many students developed more ideas for when and where to focus their attention in order to understand the role of data and what possible actions could be (as an alternative to passive use). These results reflect ideas of using the explanatory model as a lens to become informed.

According to the SEM (see Figure 5), the intention to engage with the role of data in data-driven technologies significantly affects these empowered behavior scores ($\beta = 0.164, p = 0.003$). In addition, understanding the explanatory model appears to affect empowered behavior significantly ($\beta = 0.265, p = 0.001$).

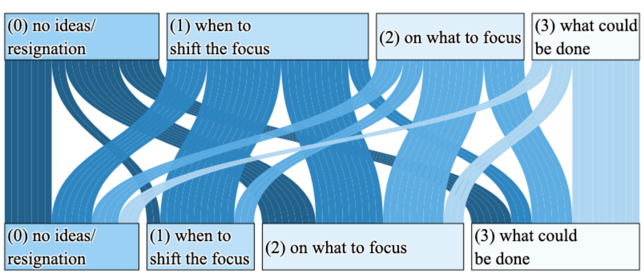

Figure 7: Students' transitions regarding empowered behavior from pre-test (top level) to post-test (bottom level)

6 DISCUSSION

In this section, we discuss the findings structured according to the research questions.

6.1 Development of Understanding of Data-Driven Technologies (RQ1)

The results showed an improved understanding of data collection, data processing purposes, and data models in data-driven technologies after the intervention. This suggests that students understood the concepts and could apply them to another everyday example. Introducing and applying the explanatory model may mitigate potential challenges of acquiring only context-specific knowledge when learning about concepts in context-based approaches [see 39, 69]. Thus, this approach appears to enhance students' comprehension of the role of data in everyday data-driven technologies, an area where students often lack understanding [e.g., 13, 36, 71, 72, 84]. While students could apply the concepts to the everyday context of the questionnaire, the explanatory model is likely to be accessible and applicable to students. Thus, it appears to aid in relating computational concepts to their everyday experiences, addressing the previously discussed challenge of such relations [e.g., 10, 33, 88].

6.2 Development of Motivations and Intentions (RQ2)

The results revealed that many students did not perceive much value in engaging with the role of data in data-driven technologies (i.e., perceiving it as useless or unnecessary) despite showing relatively high self-efficacy scores. This reluctance does not appear to stem primarily from a perceived lack of competence, contrary to explanations attributing this phenomenon to knowledge gaps [23, 44]. The struggle to grasp a value in engaging with the inner workings of such technologies may contribute to resignations [e.g., 23, 78]. Additionally, the intervention did not increase students' motivation. As the SEM showed no significant relationship between understanding and perceived value, conceptual understanding alone does not necessarily encourage students to see a value to shift the focus on the inner workings of data-driven technologies. Similar observations have been made in other studies, as learning about data practices, ML concepts, and potential impacts neither resolved feelings of powerlessness nor encouraged them to reflect critically on data-driven practices in everyday life or to change their opinions about their behavior [8, 40, 88]. This highlights the importance of considering students' motivations when examining the effects of computing education on students' everyday lives (e.g., relating concepts to everyday life) and encouraging them to perceive a personal value of using the concepts in everyday life.

In contrast to motivation, students' intentions to engage with the role of data the next time they use apps increased significantly. Since research provided evidence that implementation intentions often precede behavior initiation [35, 90], the results indicate that many students commit to engaging with the role of data in everyday interactions with such technologies. Thus, learning the explanatory model appears to encourage students to apply the model to everyday technologies and think about their interaction behavior with such technologies. This indicates that providing students with the explanatory model may reduce feelings of helplessness and powerlessness [see 18, 40, 43], suggesting a potential transition from passive roles to more informed interaction behaviors. Moreover, this underscores that students could relate the explanatory model to their everyday lives.

6.3 Development of Empowered Behavior (RQ3)

This study also explored whether learning the explanatory model could help students become informed and empowered in their everyday interactions with data-driven technologies. Given the limited research on systematically measuring empowerment, we incorporated an open-ended item to examine a sense of empowerment in students' perspectives on their behavior. We found four code categories, aligning with the goals of the explanatory model approach, which intends to support students' understanding of data-driven technologies and encourage reflections on their interactions. From this point of view, these codes could be interpreted as follows:

(0) Feeling powerless regarding the data-driven technologies and being in a passive user role
(1) Having ideas in which situations to use the explanatory model as a lens on data-driven technologies
(2) Having ideas of how to apply the explanatory model to understand data-driven technologies
(3) Evaluating specific data-driven technologies and making informed decisions for one's behavior

The question, however, is how empowerment relates to these aspects. Empowerment lacks a common definition in literature, but for instance, Freire [31] suggests empowerment involves addressing power imbalances in society to support students to understand, act, and transform the world, fostering participation in the world. Concerning computing, empowerment often entails reconstructing, understanding, and reflecting on digital technologies, alongside the capability to design digital artifacts, thereby shaping the digital world [16, 21, 49, 65, 87]. This could be interpreted with the previously discussed user-designer continuum [28, 77, 79]. In this vein, computing education aims to support transitions from passive user roles to informed and active designer roles shaping the digital world [22, 49, 80]. Taking this discussion and the findings together, students' answers revealed stages of such an empowerment at the beginning of this continuum, as illustrated in Figure 8.

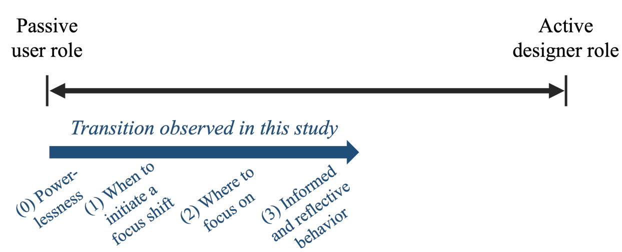

Figure 8: Findings reveal stages of becoming informed and empowered in navigating the digital world.

The results indicate decreasing powerlessness and an increase in having ideas of when and where to focus to make sense of and reflect on the role of data in data-driven technologies and thus becoming informed about such technologies. Thus, many students seem to progress toward empowerment, moving away from passive consumer stances.

This transition from powerlessness to such a form of "everyday empowerment" suggests a transformation of students' perspective on their interactions with data-driven technologies. This may result from learning the explanatory model, given the relation between understanding the explanatory model and this behavior component. Notably, students were not instructed on specific actions or behaviors in such interactions. Due to the essential role of data in data-driven technologies, peoples' stances and mindsets on the role of data can influence their readiness to engage with the inner workings of these technologies. As discussed in Section 2.4, research shows that people develop resignation and learned helplessness and feel powerless concerning data-driven practices. Moreover, learning about data-driven technologies does not necessarily help students to relate their learning to everyday life and detach themselves from such stances [see 8, 37, 40, 88]. Thus, it is challenging to support them in using their knowledge, engaging with these technologies, and reflecting on their interactions, that is, to become informed and empowered users rather than passive consumers. In this light, the study presented here provides interesting insights as learning the explanatory model appears to help students relate it to everyday life and become more empowered instead of passively using data-driven technologies in everyday life.

Future research could delve deeper into this progression of empowerment and explore the subsequent steps after being informed about the technologies and making decisions about one's behavior. For instance, informed and reflective views may open avenues for (re-) imagining future technological developments or motivate a shift to a designer role to create digital artifacts.

7 LIMITATIONS

The measurement of empowered behavior relies solely on a single open-ended item. Thus, the depth of findings about students' empowerment is limited. Nevertheless, the item allowed us to explore students' views on interacting with data-driven technologies and examine a sense of empowerment. The analysis revealed aspects that could be interpreted as incremental steps along a continuum between powerlessness and being informed and empowered. While this provides valuable insights into how understanding computational concepts could support students' empowerment when navigating the digital world, further research is needed to validate this continuum and further develop the measurement accordingly.

Additionally, the sample was a convenience sample of moderate size, geographically limited to specific regions. Although students were randomly selected through the invitation of teachers to participate with their classes, potential sample selection bias might persist, affecting the generalization of findings. However, the study was conducted in natural school settings, supporting ecological validity, and included diverse types of schools from rural and urban regions to ensure a more heterogeneous participant pool. Nonetheless, it is worth noting that attitudes and emotions towards data practices, such as resignation and powerlessness, can vary across cultures [44].

The study design focused on evaluating the presented approach as implemented in the intervention. Thus, the study did not involve a control group, precluding direct comparisons with alternative approaches. Therefore, the findings did not allow us to compare its effectiveness to other approaches. However, the study yielded valuable insights about the impact of learning the explanatory model, particularly in enhancing understanding of data-driven technologies and relating its concepts to everyday life.

8 CONCLUSIONS AND IMPLICATIONS

This paper addresses the challenge that students are struggling to relate computational concepts learned in class to their everyday lives and develop stances of resignation and powerlessness regarding the role of data in data-driven technologies. Targeting the need for respective approaches to K-12 computing education, we presented an approach called learning *explanatory models*. This approach teaches students a model about the role of data in data-driven technologies, aiming to support them in relating this to their everyday lives by providing a lens for reconstructing and evaluating data-driven technologies in everyday life. In doing so, we aim that this approach supports students to become informed and empowered instead of being resigned, powerless, and passive consumers, that is, to support their empowerment in navigating the digital world. We presented a study that evaluates this approach.

This study suggests improvements in students' conceptual understanding of data-driven technologies. Additionally, it indicates that learning the explanatory model encourages students to engage with the inner workings of everyday data-driven technologies and to reflect on their interactions with such technologies, indicating that they could relate it to their everyday experiences. Most intriguingly, the findings indicate that learning the model supports students in becoming more informed and empowered in their daily interactions with such technologies. Previous research suggested that understanding data-driven technologies does not necessarily mitigate peoples' feelings of helplessness and powerlessness. This study demonstrates how to mitigate such preventing attitudes regarding the special case of *data-driven* technologies and how transformations to informed and empowered roles in interactions with these technologies can be supported.

In addition to this approach and providing a concrete explanatory model of data-driven technologies, this paper demonstrates a way to examine students' understanding of such a model and whether they can relate this model to their everyday experiences and make use of it to become more informed and empowered in everyday life.

Moreover, the study revealed a more nuanced understanding of the stages of empowerment between passive consumer and active designer roles. This refers to a shift in focus from immediate user goals to the inner workings of these technologies and transformation in perspectives, leading to a more informed behavior in these interactions. It would be fruitful to examine further steps of this continuum, particularly regarding the transition between being informed and designing digital artifacts. Nevertheless, this continuum contributes to the ongoing debate on the potential of computing education to empower students in navigating a world shaped by data-driven technologies.

ACKNOWLEDGMENTS

This work was created as part of the ProDaBi project (www.prodabi.de), which is funded by the Deutsche Telekom Stiftung.

REFERENCES

[1] Anja Achtziger and Peter M. Gollwitzer. 2018. Motivation and Volition in the Course of Action. In *Motivation and Action*, Jutta Heckhausen and Heinz Heckhausen (Eds.). Springer International Publishing, Cham, 485–527. https://doi.org/10.1007/978-3-319-65094-4_12

[2] Icek Ajzen, Cornelia Czasch, and Michael G. Flood. 2009. From Intentions to Behavior: Implementation Intention, Commitment, and Conscientiousness. *Journal of Applied Social Psychology* 39, 6 (2009), 1356–1372.

[3] Albert Bandura. 1997. *Self-Efficacy: The Exercise of Control.* W.H. Freeman, New York.

[4] Albert Bandura. 2006. Guide for Constructing Self-Efficacy Scales. In *Self-Efficacy Beliefs of Adolescents*, Tim Urdan and Frank Pajares (Eds.). IAP, 307–338.

[5] Susan B. Barnes. 2006. A Privacy Paradox: Social Networking in the United States. *First Monday* 11, 9 (2006). https://doi.org/10.5210/fm.v11i9.1394

[6] Susanne Barth and Menno D.T. De Jong. 2017. The Privacy Paradox – Investigating Discrepancies between Expressed Privacy Concerns and Actual Online Behavior – A Systematic Literature Review. *Telematics and Informatics* 34, 7 (2017), 1038–1058. https://doi.org/10.1016/j.tele.2017.04.013

[7] Judith Bennett, Fred Lubben, and Sylvia Hogarth. 2007. Bringing Science to Life: A Synthesis of the Research Evidence on the Effects of Context-based and STS Approaches to Science Teaching. *Science Education* 91, 3 (2007), 347–370. https://doi.org/10.1002/sce.20186

[8] Karl-Emil Kjær Bilstrup, Magnus Høholt Kaspersen, Mille Skovhus Lunding, et al. 2022. Supporting Critical Data Literacy in K-9 Education: Three Principles for Enriching Pupils' Relationship to Data. In *Proceedings of the 21st Annual ACM Interaction Design and Children Conference (IDC '22)*. ACM, New York, NY, USA, 225–236. https://doi.org/10.1145/3501712.3530783

[9] Matthias Bode and Dorthe Brogaard Kristensen. 2016. The Digital Doppelgänger within: A Study on Self-Tracking and the Quantified Self Movement. In *Assembling Consumption: Researching Actors, Networks and Markets*, Robin Canniford and Domen Bajde (Eds.). Routledge, Oxon, United Kingdom, 119–135.

[10] Leanne Bowler, Amelia Acker, Wei Jeng, and Yu Chi. 2017. "It Lives All around Us": Aspects of Data Literacy in Teen's Lives. *Proceedings of the Association for Information Science and Technology* 54, 1 (2017), 27–35. https://doi.org/10.1002/pra2.2017.14505401004

[11] Robert L. Brennan and Dale J. Prediger. 1981. Coefficient Kappa: Some Uses, Misuses, and Alternatives. *Educational and Psychological Measurement* 41, 3 (1981), 687–699. https://doi.org/10.1177/001316448104100307

[12] Peter Brusilovsky and Eva Millán. 2007. User Models for Adaptive Hypermedia and Adaptive Educational Systems. In *The Adaptive Web*, Peter Brusilovsky, Alfred Kobsa, and Wolfgang Nejdl (Eds.). Vol. 4321. Springer Berlin Heidelberg, Berlin, Heidelberg, 3–53. https://doi.org/10.1007/978-3-540-72079-9_1

[13] Taina Bucher. 2017. The Algorithmic Imaginary: Exploring the Ordinary Affects of Facebook Algorithms. *Information, Communication & Society* 20, 1 (2017), 30–44. https://doi.org/10.1080/1369118X.2016.1154086

[14] Jenna Burrell. 2016. How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms. *Big Data & Society* 3, 1 (2016), 1–12. https://doi.org/10.1177/2053951715622512

[15] Lorena Casal-Otero, Alejandro Catala, Carmen Fernández-Morante, Maria Taboada, Beatriz Cebreiro, and Senén Barro. 2023. AI Literacy in K-12: A Systematic Literature Review. *International Journal of STEM Education* 10, 1 (2023), 29. https://doi.org/10.1186/s40594-023-00418-7

[16] Michael E. Caspersen, Ira Diethelm, Judith Gal-Ezer, Andrew McGettrick, Enrico Nardelli, Don Passey, Branislav Rovan, and Mary Webb. 2022. *Informatics Reference Framework for School.* NSF. https://doi.org/10.1145/3592625

[17] Center For Self-Determination Theory. [n. d.]. Intrinsic Motivation Inventory (IMI). https://selfdeterminationtheory.org/intrinsic-motivation-inventory/ accessed: 2022-09-01.

[18] Hichang Cho. 2022. Privacy Helplessness on Social Media: Its Constituents, Antecedents and Consequences. *Internet Research* 32, 1 (2022), 150–171. https://doi.org/10.1108/INTR-05-2020-0269

[19] Hanbyul Choi, Jonghwa Park, and Yoonhyuk Jung. 2018. The Role of Privacy Fatigue in Online Privacy Behavior. *Computers in Human Behavior* 81 (2018), 42–51. https://doi.org/10.1016/j.chb.2017.12.001

[20] Jacob Cohen. 1992. A Power Primer. *Psychological Bulletin* 112, 1 (1992), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

[21] Christian Dindler, Ole Sejer Iversen, Michael E. Caspersen, and Rachel Charlotte Smith. 2022. Computational Empowerment. In *Computational Thinking Education in K–12*, Siu-Cheung Kong and Harold Abelson (Eds.). The MIT Press, 121–140.

[22] Daniella DiPaola, Blakeley H. Payne, and Cynthia Breazeal. 2022. Preparing Children to Be Conscientious Consumers and Designers of AI Technologies. In *Computational Thinking Education in K–12*, Siu-Cheung Kong and Harold Abelson (Eds.). The MIT Press, 181–206. https://doi.org/10.7551/mitpress/13375.003.0014

[23] Liz Dowthwaite, Helen Creswick, Virginia Portillo, et al. 2020. "It's Your Private Information. It's Your Life.": Young People's Views of Personal Data Use by Online Technologies. In *Proceedings of the Interaction Design and Children Conference.* ACM, London United Kingdom. https://doi.org/10.1145/3392063.3394410

[24] Nora A Draper and Joseph Turow. 2019. The Corporate Cultivation of Digital Resignation. *New Media & Society* 21, 8 (2019), 1824–1839. https://doi.org/10.1177/1461444819833331

[25] Stefania Druga and Amy J Ko. 2021. How Do Children's Perceptions of Machine Intelligence Change When Training and Coding Smart Programs?. In *Interaction Design and Children.* ACM, Athens Greece, 49–61. https://doi.org/10.1145/3459990.3460712

[26] Reinders Duit, Harald Gropengießer, Ulrich Kattmann, Michael Komorek, and Ilka Parchmann. 2012. The Model of Educational Reconstruction – a Framework for Improving Teaching and Learning Science1. In *Science Education Research and Practice in Europe: Retrosspective and Prospecctive*, Doris Jorde and Justin Dillon (Eds.). SensePublishers, Rotterdam, 13–37. https://doi.org/10.1007/978-94-6091-900-8_2

[27] Sally Fincher, Johan Jeuring, Craig S. Miller, Peter Donaldson, Benedict Du Boulay, Matthias Hauswirth, Arto Hellas, Felienne Hermans, Colleen Lewis, Andreas Mühling, Janice L. Pearce, and Andrew Petersen. 2020. Notional Machines in Computing Education: The Education of Attention. In *Proceedings of the Working Group Reports on Innovation and Technology in Computer Science Education*. ACM, Trondheim Norway, 21–50. https://doi.org/10.1145/3437800.3439202

[28] Gerhard Fischer and Elisa Giaccardi. 2006. Meta-Design: A Framework for the Future of End-User Development. In *End User Development*, Henry Lieberman, Fabio Paternò, and Volker Wulf (Eds.). Vol. 9. Springer Netherlands, Dordrecht, 427–457. https://doi.org/10.1007/1-4020-5386-X_19

[29] Jessica Fishman, Viktor Lushin, and David S. Mandell. 2020. Predicting Implementation: Comparing Validated Measures of Intention and Assessing the Role of Motivation When Designing Behavioral Interventions. *Implementation Science Communications* 1, 1 (2020). https://doi.org/10.1186/s43058-020-00050-4

[30] Jessica Flake, Kenneth Barron, Chris Hulleman, D. Betsy Mccoach, and Megan Welsh. 2015. Measuring Cost: The Forgotten Component of Expectancy-Value Theory. *Contemporary Educational Psychology* 41 (2015). https://doi.org/10.1016/j.cedpsych.2015.03.002

[31] Paulo Freire. 2012. *Pedagogy of the Oppressed* (repr ed.). Bloomsbury, New York.

[32] Kamel Gana and Guillaume Broc. 2019. *Structural Equation Modeling with Lavaan*. ISTE, London.

[33] Engida H. Gebre. 2018. Young Adults' Understanding and Use of Data: Insights for Fostering Secondary School Students' Data Literacy. *Canadian Journal of Science, Mathematics and Technology Education* 18, 4 (2018), 330–341. https://doi.org/10.1007/s42330-018-0034-z

[34] John K. Gilbert. 2006. On the Nature of "Context" in Chemical Education. *International Journal of Science Education* 28, 9 (2006), 957–976. https://doi.org/10.1080/09500690600702470

[35] Peter Gollwitzer and Paschal Sheeran. 2006. Implementation Intentions and Goal Achievement: A Meta-analysis of Effects and Processes. *Advances in Experimental Social Psychology* 38 (2006). https://doi.org/10.1016/S0065-2601(06)38002-1

[36] Cami Goray and Sarita Schoenebeck. 2022. Youths' Perceptions of Data Collection in Online Advertising and Social Media. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–27. https://doi.org/10.1145/3555576

[37] Gregor Große-Bölting and Andreas Mühling. 2020. Students Perception of the Inner Workings of Learning Machines. In *2020 International Conference on Learning and Teaching in Computing and Engineering (LaTICE)*, Vol. 5. IEEE Computer Society, Ho Chi Minh City.

[38] Shuchi Grover. 2024. Teaching AI to K-12 Learners: Lessons, Issues, and Guidance. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1*. ACM, Portland OR USA, 422–428. https://doi.org/10.1145/3626252.3630937

[39] Mark Guzdial. 2010. Does Contextualized Computing Education Help? *ACM Inroads* 1, 4 (2010), 4–6. https://doi.org/10.1145/1869746.1869747

[40] Eszter Hargittai and Alice Marwick. 2016. "What Can I Really Do?" Explaining the Privacy Paradox with Online Apathy. *Int. Journal of Communication* 10 (2016), 3737–3757.

[41] Jutta Heckhausen. 2007. The Motivation-Volition Divide and Its Resolution in Action-Phase Models of Developmental Regulation. *Research in Human Development* 4, 3-4 (2007), 163–180. https://doi.org/10.1080/15427600701662983

[42] Tom Hitron, Yoav Orlev, Iddo Wald, Ariel Shamir, Hadas Erel, and Oren Zuckerman. 2019. Can Children Understand Machine Learning Concepts?: The Effect of Uncovering Black Boxes. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland Uk, 1–11. https://doi.org/10.1145/3290605.3300645

[43] Christian Pieter Hoffmann, Christoph Lutz, and Giulia Ranzini. 2016. Privacy Cynicism: A New Approach to the Privacy Paradox. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace* 10, 4 (2016). https://doi.org/10.5817/CP2016-4-7

[44] Christian Pieter Hoffmann, Christoph Lutz, and Giulia Ranzini. 2024. Inequalities in Privacy Cynicism: An Intersectional Analysis of Agency Constraints. *Big Data & Society* 11, 1 (2024). https://doi.org/10.1177/20539517241232629

[45] Lukas Höper and Carsten Schulte. 2024. The Data Awareness Framework as Part of Data Literacies in K-12 Education. *Information and Learning Sciences* 125, 7/8 (2024), 491–512. https://doi.org/10.1108/ILS-06-2023-0075

[46] Lukas Höper and Carsten Schulte. 2024. Empowering Students for the Data-Driven World: A Qualitative Study of the Relevance of Learning about Data-Driven Technologies. *Informatics in Education* (2024). https://doi.org/10.15388/infedu.2024.19

[47] Li-tze Hu and Peter M. Bentler. 1999. Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives. *Structural Equation Modeling: A Multidisciplinary Journal* 6, 1 (1999), 1–55. https://doi.org/10.1080/10705519909540118

[48] Ole Sejer Iversen, Rachel Charlotte Smith, and Christian Dindler. 2018. From Computational Thinking to Computational Empowerment: A 21$^{\text{st}}$ Century PD Agenda. In *Proceedings of the 15th Participatory Design Conference: Full Papers - Volume 1*. ACM, Hasselt and Genk Belgium, 1–11. https://doi.org/10.1145/3210586.3210592

[49] K–12 Computer Science Framework Steering Committee. 2016. K-12 Computer Science Framework. http://www.k12cs.org

[50] Caroline Keen. 2020. Apathy, Convenience or Irrelevance? Identifying Conceptual Barriers to Safeguarding Children's Data Privacy. *New Media & Society* 24, 1 (2020), 1–20. https://doi.org/10.1177/1461444820960068

[51] Rob Kitchin. 2014. *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences*. SAGE Publications, Los Angeles, California.

[52] A. D. I. Kramer, J. E. Guillory, and J. T. Hancock. 2014. Experimental Evidence of Massive-Scale Emotional Contagion through Social Networks. *Proceedings of the National Academy of Sciences* 111, 24 (2014), 8788–8790. https://doi.org/10.1073/pnas.1412583111

[53] Udo Kuckartz. 2014. *Qualitative Text Analysis: A Guide to Methods, Practice & Using Software*. SAGE, Los Angeles.

[54] J. Richard Landis and Gary G. Koch. 1977. The Measurement of Observer Agreement for Categorical Data. *Biometrics* 33, 1 (1977), 159–174. http://www.jstor.org/stable/2529310

[55] Irene Lee, Safinah Ali, Helen Zhang, Daniella DiPaola, and Cynthia Breazeal. 2021. Developing Middle School Students' AI Literacy. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*. ACM, Virtual Event USA, 191–197. https://doi.org/10.1145/3408877.3432513

[56] Sonia Livingstone, Mariya Stoilova, and Rishita Nandagiri. 2019. Children's Data and Privacy Online: Growing up in a Digital Age. An Evidence Review.

[57] Duri Long and Brian Magerko. 2020. What Is AI Literacy? Competencies and Design Considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–16. https://doi.org/10.1145/3313831.3376727

[58] Steven F. Maier and Martin E. Seligman. 1976. Learned Helplessness: Theory and Evidence. *Journal of Experimental Psychology: General* 105, 1 (1976), 3–46. https://doi.org/10.1037/0096-3445.105.1.3

[59] David Markland and Lew Hardy. 1997. On the Factorial and Construct Validity of the Intrinsic Motivation Inventory: Conceptual and Operational Concerns. *Research Quarterly for Exercise and Sport* 68, 1 (1997), 20–32. https://doi.org/10.1080/02701367.1997.10608863

[60] Karl Maton. 2013. Making Semantic Waves: A Key to Cumulative Knowledge-Building. *Linguistics and Education* 24, 1 (2013), 8–22. https://doi.org/10.1016/j.linged.2012.11.005

[61] Edward McAuley, Terry Duncan, and Vance V. Tammen. 1989. Psychometric Properties of the Intrinsic Motivation Inventory in a Competitive Sport Setting: A Confirmatory Factor Analysis. *Research Quarterly for Exercise and Sport* 60, 1 (1989), 48–58. https://doi.org/10.1080/02701367.1989.10607413

[62] Pekka Mertala and Janne Fagerlund. 2024. Finnish 5th and 6th Graders' Misconceptions about Artificial Intelligence. *International Journal of Child-Computer Interaction* 39 (2024). https://doi.org/10.1016/j.ijcci.2023.100630

[63] Pekka Mertala, Janne Fagerlund, and Oscar Calderon. 2022. Finnish 5th and 6th Grade Students' Pre-Instructional Conceptions of Artificial Intelligence (AI) and Their Implications for AI Literacy Education. *Computers and Education: Artificial Intelligence* 3 (2022). https://doi.org/10.1016/j.caeai.2022.100095

[64] Joseph E Michaelis and David Weintrop. 2022. Interest Development Theory in Computing Education: A Framework and Toolkit for Researchers and Designers. *ACM Transactions on Computing Education* 22, 4 (2022), 1–27. https://doi.org/10.1145/3487054

[65] Luis Morales-Navarro and Yasmin B. Kafai. 2023. Conceptualizing Approaches to Critical Computing Education: Inquiry, Design, and Reimagination. In *Past, Present and Future of Computing Education Research*, Mikko Apiola, Sonsoles López-Pernas, and Mohammed Saqr (Eds.). Springer International Publishing, Cham, 521–538. https://doi.org/10.1007/978-3-031-25336-2_21

[66] Rainer Mühlhoff. 2021. Predictive Privacy: Towards an Applied Ethics of Data Analytics. *Ethics and Information Technology* 23, 4 (2021), 675–690. https://doi.org/10.1007/s10676-021-09606-x

[67] Andreas Mühling and Gregor Große-Bölting. 2023. Novices' Conceptions of Machine Learning. *Computers and Education: Artificial Intelligence* 4 (2023). https://doi.org/10.1016/j.caeai.2023.100142

[68] Davy Tsz Kit Ng, Jac Ka Lok Leung, Samuel Kai Wah Chu, and Maggie Shen Qiao. 2021. Conceptualizing AI Literacy: An Exploratory Review. *Computers and Education: Artificial Intelligence* 2 (2021). https://doi.org/10.1016/j.caeai.2021.100041

[69] Jacqueline Nijenhuis-Voogt, Durdane Bayram-Jacobs, Paulien C. Meijer, and Erik Barendsen. 2021. Omnipresent yet Elusive: Teachers' Views on Contexts for Teaching Algorithms in Secondary Education. *Computer Science Education* 31, 1 (2021), 30–59. https://doi.org/10.1080/08993408.2020.1783149

[70] OECD. 2014. Summary of the OECD Privacy Expert Roundtable: Protecting Privacy in a Data-driven Economy: Taking Stock of Current Thinking. https://doi.org/10.1787/9789264196391-en

[71] Luci Pangrazio and Neil Selwyn. 2019. 'Personal Data Literacies': A Critical Literacies Approach to Enhancing Understandings of Personal Digital Data. *New Media & Society* 21, 2 (2019), 419–437. https://doi.org/10.1177/1461444818799523

[72] Luci Pangrazio and Neil Selwyn. 2020. Towards a School-Based 'Critical Data Education'. *Pedagogy, Culture & Society* 29, 3 (2020), 431–448. https://doi.org/10.1080/14681366.2020.1747527

[73] Yim Register and Amy J. Ko. 2020. Learning Machine Learning with Personal Data Helps Stakeholders Ground Advocacy Arguments in Model Mechanics. In *Proceedings of the 2020 ACM Conference on International Computing Education Research*. ACM, Virtual Event New Zealand, 67–78. https://doi.org/10.1145/3372782.3406252

[74] Mitchel Resnick and Brian Silverman. 2005. Some Reflections on Designing Construction Kits for Kids. In *Proceedings of the 2005 Conference on Interaction Design and Children*. ACM, Boulder Colorado, 117–122. https://doi.org/10.1145/1109540.1109556

[75] Saman Rizvi, Jane Waite, and Sue Sentence. 2023. Artificial Intelligence Teaching and Learning in K-12 from 2019 to 2022: A Systematic Literature Review. *Computers and Education: Artificial Intelligence* 4 (2023). https://doi.org/10.1016/j.caeai.2023.100145

[76] Michael T. Rücker, Wouter R. Van Joolingen, and Niels Pinkwart. 2020. Small but Powerful: A Learning Study to Address Secondary Students' Conceptions of Everyday Computing Technology. *ACM Transactions on Computing Education* 20, 2 (2020), 1–27. https://doi.org/10.1145/3377880

[77] Douglas Rushkoff. 2010. *Program or Be Programmed: Ten Commands for a Digital Age*. OR Books, New York.

[78] Ina Sander. 2020. What Is Critical Big Data Literacy and How Can It Be Implemented? *Internet Policy Review* 9, 2 (2020). https://doi.org/10.14763/2020.2.1479

[79] Carsten Schulte and Lea Budde. 2018. A Framework for Computing Education: Hybrid Interaction System: The Need for a Bigger Picture in Computing Education. In *Proceedings of the 18th Koli Calling International Conference on Computing Education Research*. ACM, Koli Finland, 1–10. https://doi.org/10.1145/3279720.3279733

[80] Carsten Schulte and Maria Knobelsdorf. 2007. Attitudes towards Computer Science-Computing Experiences as a Starting Point and Barrier to Computer Science. In *Proceedings of the Third International Workshop on Computing Education Research - ICER '07*. ACM Press, Atlanta, Georgia, USA, 27. https://doi.org/10.1145/1288580.1288585

[81] Dale H. Schunk, Maria K. DiBenedetoo, Kathryn R. Wentzel, and David B. Miele. 2016. Self-Efficacy Theory in Education. In *Handbook of Motivation at School* (2 ed.). Routledge, New York, 34–54.

[82] D Sculley, Gary Holt, Daniel Golovin, Eugene Davydov, Todd Phillips, Dietmar Ebner, Vinay Chaudhary, Michael Young, Jean-François Crespo, and Dan Dennison. 2015. Hidden Technical Debt in Machine Learning Systems. In *NIPS'15: Proceedings of the 28th International Conference on Neural Information Processing Systems*, Vol. 2. 2503–2511.

[83] Daniel Susser, Beate Roessler, and Helen Nissenbaum. 2019. Technology, Autonomy, and Manipulation. *Internet Policy Review* 8, 2 (2019), 1–22. https://doi.org/10.14763/2019.2.1410

[84] Matti Tedre, Henriikka Vartiainen, Juho Kahila, Tapani Toivonen, Ilkka Jormanainen, and Teemu Valtonen. 2020. Machine Learning Introduces New Perspectives to Data Agency in K—12 Computing Education. In *2020 IEEE Frontiers in Education Conference (FIE)*. IEEE, Uppsala, 1–8. https://doi.org/10.1109/FIE44824.2020.9274138

[85] David Touretzky, Christina Gardner-McCune, Fred Martin, and Deborah Seehorn. 2019. Envisioning AI for K-12: What Should Every Child Know about AI? *Proceedings of the AAAI Conference on Artificial Intelligence* 33 (2019), 9795–9799. https://doi.org/10.1609/aaai.v33i01.33019795

[86] Zeynep Tufekci. 2014. Engineering the Public: Big Data, Surveillance and Computational Politics. *First Monday* 19, 7 (2014). https://doi.org/10.5210/fm.v19i7.4901

[87] Maarten Van Mechelen, Line Have Musaeus, Ole Sejer Iversen, Christian Dindler, and Arthur Hjorth. 2021. A Systematic Review of Empowerment in Child-Computer Interaction Research. In *Interaction Design and Children*. ACM, Athens Greece, 119–130. https://doi.org/10.1145/3459990.3460701

[88] Henriikka Vartiainen, Tapani Toivonen, Ilkka Jormanainen, Juho Kahila, Matti Tedre, and Teemu Valtonen. 2021. Machine Learning for Middle Schoolers: Learning through Data-Driven Design. *International Journal of Child-Computer Interaction* 29 (2021), 1–12. https://doi.org/10.1016/j.ijcci.2021.100281

[89] Jane Waite, Karl Maton, Paul Curzon, and Lucinda Tuttiett. 2019. Unplugged Computing and Semantic Waves: Analysing Crazy Characters. In *Proceedings of the 1st UK & Ireland Computing Education Research Conference on - UKICER*. ACM Press, Canterbury, United Kingdom, 1–7. https://doi.org/10.1145/3351287.3351291

[90] Thomas L. Webb and Paschal Sheeran. 2006. Does Changing Behavioral Intentions Engender Behavior Change? A Meta-Analysis of the Experimental Evidence. *Psychological Bulletin* 132, 2 (2006), 249–268.

[91] Allan Wigfield and Jenna Cambria. 2010. Students' Achievement Values, Goal Orientations, and Interest: Definitions, Development, and Relations to Achievement Outcomes. *Developmental Review* 30, 1 (2010), 1–35. https://doi.org/10.1016/j.dr.2009.12.001

[92] Allan Wigfield and Jacquelynne S. Eccles. 2000. Expectancy–Value Theory of Achievement Motivation. *Contemporary Educational Psychology* 25, 1 (2000), 68–81. https://doi.org/10.1006/ceps.1999.1015

[93] Allan Wigfield and Jacquelynne S. Eccles. 2023. *The Relevance of Situated Expectancy-Value Theory to Understanding Motivation and Emotion in Different Contexts* (1 ed.). Routledge, London, 3–18.

[94] Abigail Zimmermann-Niefield, Makenna Turner, Bridget Murphy, Shaun K. Kane, and R. Benjamin Shapiro. 2019. Youth Learning Machine Learning through Building Models of Athletic Moves. In *Proceedings of the 18th ACM International Conference on Interaction Design and Children*. ACM, Boise ID USA, 121–132. https://doi.org/10.1145/3311927.3323139

[95] Shoshana Zuboff. 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (first edition ed.). PublicAffairs, New York.