

Measuring the Data Agency of Pre-Service Teachers: A Six-Factor Model

Teemu VALTONEN¹ [0000-0002-1803-9865], Henriikka VARTIAINEN¹ [0000-0001-6005-907X], Eliisa VÄHÄ¹ [0009-0007-0788-9518], Jari KUKKONEN¹ [0000-0002-9448-4784], Erkki SOINTU² [0000-0003-4001-7264], Juho KAHILA¹ [0000-0002-9913-0627], Matti TEDRE³ [0000-0003-1037-3313]

¹University of Eastern Finland, Faculty of Philosophy, School of Applied Educational Science and Teacher Education

²University of Eastern Finland, Faculty of Philosophy, School of Educational Sciences and Psychology

³University of Eastern Finland, Faculty of Science, Forestry and Technology, School of Computing

E-mail addresses: teemu.valtonen@uef.fi; henriikka.vartiainen@uef.fi; eliisa.vaha@uef.fi; jari.kukkonen@uef.fi; erikko.sointu@uef.fi; juho.kahila@uef.fi; matti.tedre@uef.fi

Abstract. This article examines pre-service teachers' data agency, defined as the ability to act according to one's own values and goals rather than being directed by algorithmic systems. Data agency involves understanding how computational systems, such as algorithms, data-driven profiling, and platform infrastructures, collect, process, and use data, and how these practices shape individuals and society. This article introduces a self-assessment instrument developed to measure data agency and applies it to a sample of 163 Finnish pre-service teachers. The findings show that pre-service teachers evaluated their competencies across different dimensions of data agency rather cautiously. The study highlights the importance of strengthening future teachers' understanding of the mechanisms behind algorithmic and data-driven decision-making. Such knowledge is increasingly essential for preparing future teachers to address challenges related to datafication, including commercial data collection and algorithmic influencing in contemporary education.

Keywords: data agency, datafication, pre-service teacher, survey, social media literacy, AI literacy, AI education.

1. Introduction

Before the 2010s, it was common to say that the age of algorithms would require people to acquire new programming-oriented mindsets, such as algorithmic thinking (Harel, 2004), procedural thinking (Papert, 1980) or computational thinking (Wing, 2006). Many types of literacy were suggested as essential additions to the K-12 curriculum, such as proce-

dural literacy (Sheil, 1980), computational literacy (diSessa, 2000), literacy in algorithmic reasoning (Culbertson, 1986), procedurality (Vee, 2010) and fluency with information technology, or IT (NRC Committee on Information Technology Literacy, 1999; see also Tedre and Denning, 2016). On the emerging horizon of new societal challenges, research pointed out how information systems shape various aspects of life, including work (Zuboff, 1988), worldviews (Bowker & Star, 2000), identity formation (boyd, 2015) and social bonds (boyd & Ellison, 2007). Discussions of concepts like information literacy, media literacy and digital literacy, among many others (Bawden, 2001), emphasised the knowledge and skills necessary in a looming information society (Castells, 1996).

As the datafication of society progressed, attention shifted from data-processing tools (e.g. the computer) and data-transforming processes (the algorithm) to data itself (Hintz et al., 2018; Zuboff, 2019). Advances in pervasive computing, data-intensive analysis, cloud computing, social media, and the Internet of Things facilitated unparalleled tracking and profiling capabilities and offered granular insights into individual behaviours (Swan, 2012; Zuboff, 2019). While developments in collecting and analysing sensitive data at a massive scale have been noteworthy, the actual achievement lies in automating decision-making using this data and transforming raw data into actionable insights, often at the user's expense. Concepts like data literacy and information privacy were proposed as essential knowledge and skills for understanding the processes of amassing substantial data – covering locations, moods, preferences, and even personal secrets – as individuals interacted with everyday digital platforms (Bowler et al., 2017; Swan, 2012). A massive, interdisciplinary body of research on data and information privacy contributed to understanding the gradual loss of privacy from perspectives of rights, commodities, and control (Smith et al., 2011).

At the same time, as nuances of data literacy gained recognition across various domains (Bowler et al., 2017; Cui et al., 2023; Hintz et al., 2018; Ridsdale et al., 2015), an increasing number of scholars advocated for a more profound transition towards data-related agency (Tedre et al., 2020; Kennedy et al., 2015; Lehtiniemi & Haapoja, 2020; Lupton & Williamson, 2017). Data agency refers to people's volition and capacity to take informed actions that make a difference in their digital world (Vartiainen et al., 2021). It encompasses the ability to regulate what data traces one leaves, affect their processing and their consequences, weigh and offset trade-offs in using data-driven services, identify and work out associated ethical, legal and societal problems and engage in online activities without risking safety and security (Vartiainen et al., 2021; Kennedy et al., 2015). Data agency has been linked with the ability to control information flows, manipulate them and use them wisely (Selwyn and Pangrazio, 2018) and especially to resist behaviour engineering, targeted advertisements and manipulative influencing (Valtonen et al., 2019).

Data agency has also become an important topic in education (Valtonen et al., 2019). The datafication of society and the adoption of social media have placed even young children in the middle of datafied society (boyd, 2015; boyd & Marwick, 2011; Davis & James, 2013; Hargittai & Marwick, 2016; Lapenta & Jørgensen, 2015; Lievens et al., 2019; Raynes-Goldie, 2010; Selwyn & Pangrazio, 2018; Stoilova et al., 2019). Understanding and learning about mechanisms of datafication has become crucially important for children begin-

ning to develop their data agency, related security mindsets and personal data strategies (Lupton & Williamson, 2017). These themes pose new areas and challenges for schools, teachers, and teacher training. In alignment with the core curriculum of Finland (Finnish National Board of Education, 2014), students need to be active participants in a democratic society, able to think critically about and contribute to shared societal issues, as well as to areas where datafication and active agency play important roles. Again, this poses new expectations for teachers and teacher training and questions about teachers' readiness and data agency.

From the perspective of computing education, data agency is not a civic competence or a media literacy competence, but essentially a computational one (Berland & Garcia, 2024). Understanding how data are generated, collected, clustered, analyzed, and used in systems requires familiarity with core computing concepts, such as algorithms, data models, recommender systems, and automated decision-making (Höper & Schulte, 2024). A lack of such understanding limits learners' ability to exercise agency in datafied environments (Höper & Schulte, 2024). As international and national curricula increasingly address topics related to AI, platformization, and algorithms (e.g., Miao, 2022; Miao & Shiohira, 2024), teachers' own data agency becomes a key prerequisite for meaningful computing education.

Although little is known about pre-service teachers' experiences, they play a crucial role in translating data literacy and agency into educational practice (Vartiainen et al., 2021). Accordingly, this research examines data agency from the perspective of pre-service teachers by introducing an instrument to measure it. Many assessment instruments exist for concepts such as data literacy and information privacy (see reviews by Cui et al., 2023; Smith et al., 2011), but for datafication and data agency research, a more focused and contextually relevant instrument is needed. This article presents a self-evaluation instrument grounded in key literature on datafication and data agency and reports perceptions of data agency among 163 Finnish pre-service teachers.

2. Theoretical framework

Growing amounts of data about people's everyday activities are harvested by commercial digital platform providers (Kitchin, 2012). Combined with AI, these data influence and affect people's decisions and choices, typically without their realisation (e.g. Hendricks & Vestergaard, 2018; Valtonen et al., 2019; Zuboff, 2015). This development undermines individuals' agency as they are unaware of this influence on their actions.

The concept of agency originates from sociology and psychology (see Bandura, 2001; Emirbayer & Mische, 1998). From the perspective of social cognitive theory, agency is understood as intentionality, including action plans and regulatory strategies for implementing them (Bandura, 2001). From the perspective of sociocultural theories, agency is seen as goal-driven activities and decision-making mediated by certain sociocultural, technological and material circumstances that shape or constrain human agency (Wertsch, 2007). Data agency goes beyond data and information literacy by highlighting the need

for individual volition, autonomy and capability for informed actions (Tedre et al., 2020; Vartiainen et al., 2021; Lehtiniemi & Haapoja, 2020).

Information and data literacies focus on the ability to locate, evaluate, interpret, and use information and data effectively. Data agency in turn goes beyond comprehension emphasizing one's control, decision-making, and ethical engagement in data-intensive environments. Data agency refers to individuals' abilities to make their own acknowledged choices and decisions, while being influenced by algorithmic systems and data-driven platforms. It emphasises awareness of the role of data collected from users as part of personalised commercial online environments and media, influencing choices and decision-making, strategies and tactics (Kennedy et al., 2015), power and participation (Selwyn & Pangrazio, 2018; Wenger, 2023) and active control of personal data (Lehtiniemi & Haapoja, 2020). Data agency can be seen as a multidimensional entity, demanding that individuals have the understanding, skills, and attitudes necessary to influence and adjust the ability of different online platforms to collect and use the data they leave behind (Vartiainen et al., 2021; Kennedy et al., 2015; Lehtiniemi & Haapoja, 2020; Selwyn & Pangrazio, 2018). This article focuses on three elements (Figure 1) important for agency (Bandura, 2001, 2018; Emirbayer & Mische, 1998) in the context of the digital world: understanding and awareness, skills and capabilities, and attitudes.

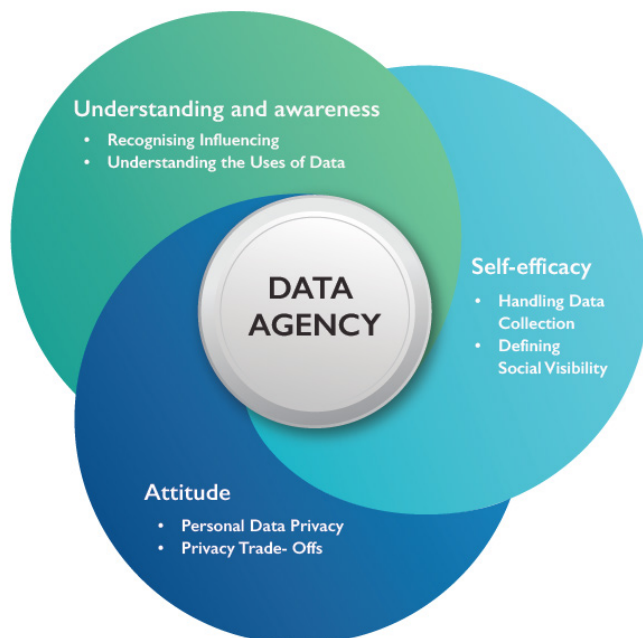


Figure 1: Three elements of the Data Agency.

2.1 *Element 1: Understanding and awareness*

Understanding some key dynamics and mechanisms of social media and other key players in the datafication of society is a rationale behind many initiatives for personal data literacy, AI literacy and many other new literacies (e.g. Höper & Schulte, 2023). People often provide their data unconsciously and with little understanding of the value of those data and their possible uses (Tedre et al. 2020), while platform providers see data as a highly valuable asset for commercial and political actors (van Dijck, 2013). There is evidence that youth may unwittingly provide personal or even sensitive data without comprehending the extent of its collection and use (Tedre et al., 2020; Keen, 2020; Lapenta & Jørgensen, 2015; Selwyn & Pangrazio, 2018). Understanding and awareness of datafication raises knowledge of data collection and application in everyday life. Consider how people's online actions create data, or how those data are analysed and used for various purposes such as profiling, modelling, recommending and influencing (Bowler et al., 2017; Pangrazio & Selwyn, 2019; Valtonen et al., 2019). Understanding the lifecycle of data, the ubiquity of its collection and the ways it can be used for various purposes is a prerequisite for informed choices, decisions and, importantly, agency. It is key to recognising influencing attempts and mis/dis/mal information (Carmi et al., 2020) and discovering when and how one's data are used in large-scale and data-driven systems (Wenger, 2023).

The understanding and awareness of datafication has many dimensions. It includes, for example, understanding the price of the free applications and how personal data are turned into commodities to be sold. This is a prerequisite for informed actions and personal data strategies. It includes understanding how the data collected through free applications and entertainment enable powerful measures, such as profiling, targeted marketing, behaviour engineering, fake news and targeted disinformation (Carmi et al., 2020; Hendricks & Vestergaard, 2018; Kramer et al., 2014). Like-minded media amplify political polarisation based on similar online preferences (Kubin & von Sikorski, 2021). Understanding the logic of datafication is needed for understanding the possibilities of influencing (commercial, political, religious or other) for critical positions towards personalised content and for agency in the digital world.

2.2 *Element 2: Self-efficacy*

Studies on agency highlight self-efficacy, self-beliefs about the capacity to conduct certain activities, as a central element (Bandura, 1982). In data agency, self-efficacy refers to people's beliefs in their ability to influence their online presence, data traces, social media visibility and privacy settings. Such abilities are central to people's data strategies (Selwyn & Pangrazio, 2018). Skills that underlie data agency enable young people to employ strategies to shape their social presence on social media and the Internet (see Davis & James, 2013; Hargittai & Marwick, 2016; Tufekci, 2008). They facilitate agency and reflexivity, giving people a voice and power in the data-driven world (Couldry & Powell, 2014). They are essential for data activism and for re-articulating old notions such as democracy and participation (Baack, 2015; Lehtiniemi & Haapoja, 2020). Data-related skills enable data

activists to turn datafication from a tool of oppression into a means to achieve their goals (Baack, 2015).

Based on previous studies, it seems that youth are skilled in defining their online visibility and presence (see Davis & James, 2013; Hargittai & Marwick, 2016; Tufekci, 2008). Their tactics vary by service and social context, including friends-only settings and account-switching (boyd & Marwick, 2011; Davis & James, 2013; Hargittai & Marwick, 2016). In the latter, individuals use different platforms for different social circles and purposes: for instance, Facebook for family and social purposes, Instagram for friends and entertainment purposes and Twitter for broader social and informational engagement (Pelletier et al., 2020). In the former, youth often apply ‘friends-only’ settings to restrict the visibility of their posts to a select group, thereby controlling their data’s exposure (boyd & Marwick, 2011; Davis & James, 2013; Hargittai & Marwick, 2016; Marwick & boyd, 2014; Tufekci, 2008). Furthermore, many account-switch or use pseudonyms to compartmentalise different personas they want to maintain, such as a student persona and a hobbyist persona (Davis & James, 2013; Hargittai & Marwick, 2016). Others enter incorrect details (Davis & James, 2013), remove or block individuals (Stoilova et al., 2020), erase wall posts, comments and tags (Davis & James, 2013; Raynes-Goldie, 2010) and employ language and imagery understood only within a certain social circle (Marwick & boyd, 2014).

These acts show that youth seem to be rather well-equipped to shape their social presence and appearance online. Still, the challenge from the perspective of the data agency is that youth tend to spend more effort on managing their online visibility and social reputation among peers, friends, and family than on the risks posed by commercial tracking and data harvesting (boyd, 2015; Keen, 2020; Lapenta & Jørgensen, 2015).

2.3 Element 3: Attitude

Agency requires an individual to adopt a certain attitude towards a phenomenon or behaviour (Eteläpelto et al., 2013). Within the datafied society, appreciation and indifference toward personal data constitute the third area of data agency. Privacy- and security-oriented attitudes – especially context-dependent concerns – play a role in digital data management (Acquisti et al., 2015; Hargittai & Marwick, 2016; Selwyn & Pangrazio, 2018). Context-specific concerns and worries shape the ongoing privacy calculus of the daily trade-offs one must make between privacy and access to services (Acquisti et al., 2015). They range from context-specific and individual concerns, such as mobile media users’ informed stances towards the use of their data (which facilitate conscious forms of data agency) (Selwyn & Pangrazio, 2018), to abstract values and ideals, such as those underlying the open data movement (Baack, 2015).

Young people tend to be more concerned about social privacy over commercial tracking (boyd, 2015; Keen, 2020; Lapenta & Jørgensen, 2015; Lievens et al., 2019). The role of commercial tracking and its possibilities are not viewed as such a risk that it requires much concern (boyd, 2015; Keen, 2020; Lapenta & Jørgensen, 2015). Research on the commercial dimensions of online privacy has shown that youth take precautionary measures

to limit commercial tracking. However, these actions are often less common than what is executed on a personal level (Hargittai & Marwick, 2016). Lievens et al. (2019) explained this by the immediacy of interpersonal privacy risks compared to the long-term, often more indistinct implications of commercial data collection. On the commercial front, although some youth engage in practices like using incognito browser modes and plugins that prevent tracking or deleting cookies, these actions often occur alongside an acceptance that personalised data-based advertisements are an acceptable trade-off for free access to desired online services (Lapenta & Jørgensen, 2015; Selwyn & Pangrazio, 2018). The indifference has even been described as apathy over commercial dimensions of online privacy (Hargittai & Marwick, 2016; Selwyn & Pangrazio, 2018). Youth often perceive data privacy as the obligation of the platform providers, and end-user license agreements, typically written in confusing legal jargon (Lapenta & Jørgensen, 2015), are seldom read (Keen, 2020; Lapenta & Jørgensen, 2015; Selwyn & Pangrazio, 2018). It has been noted that young people often view platforms and service providers as non-judgmental and innocuous, leading them to lower their guard against data collection (Keen, 2020). Research has suggested that the normalisation of ubiquitous data collection across all sectors of life fuels indifference, apathy and cynical views about its unavoidability in commercial privacy (Hargittai & Marwick, 2016; Selwyn & Pangrazio, 2018).

These three areas, understanding and awareness of data, skills and capacities for active and autonomous actions, and attitudes towards collecting personal data, create the data agency framework. These areas align with traditional computing education goals of helping learners understand how algorithmic and AI-driven systems can transform data into decisions, and how these systems can be interrogated, designed, and resisted. Data agency connects classical computational thinking and data-driven thinking with ethical reasoning, critical data literacy, and participation in datafied societies. The following section describes approaches for operationalising this framework into a questionnaire capable of measuring perceptions of data agency.

3. Methodology

This research aims to develop a self-evaluation instrument for measuring pre-service teachers' data agency and their perceptions of it. It is distilled from previous literature focusing on three main elements of data agency: understanding and awareness of data, skills and capacities for active and autonomous actions, and attitudes towards the value of personal data. The instrument was designed to emphasise individuals' conceptions of their data agency. The research questions were:

1. What are the psychometric properties of the self-report instrument synthesised from the literature for measuring pre-service teachers' data agency?
2. How do pre-service teachers perceive their data agency through the three elements?
3. What kinds of differences between pre-service teachers can be identified?
4. How are data agency areas related to each other?

The results will assess the need for developing a data agency during pre-service teacher training.

3.1 Instrument design

The instrument's design was based on previous articles in the field of computing education targeting data agency (Vartiainen et al., 2024; Tedre et al., 2020) with previous definitions of the topic, including, e.g., algorithmic influencing, data collection mechanisms, and data-driven platforms, used as a starting point. The definitions were deepened with current literature about AI literacy, data ethics, and algorithmic fairness, to better conceptualise the areas of data agency. After this, the areas were further operationalised as statements measuring the three main areas. The survey instrument was developed in Finnish, which is the participants' primary language of the target group. The English translation, reported in this article, was carried out by researchers familiar with the research domain to ensure semantic and conceptual equivalence with the original Finnish items. These processes were conducted in collaboration with a team of experts from education, teacher training and computer science. The creation of statements produced a pool of items. This article describes the testing of the instrument's psychometric properties using confirmatory factor analysis (CFA). After the CFA, the final factors were formed and named (Table 1).

Table 1: Items in the data agency instrument

Category	Statements
Understanding and Awareness	In social media: <ul style="list-style-type: none"> • I often notice content that I recognise as being intentionally targeted at me. • I often come across various content that aims to influence my political views. • I often notice content that aims to influence my values. • I receive a lot of content intended to affect my actions.
	Understanding the Uses of Data (UUD) <ul style="list-style-type: none"> • I know how social media collects information about me. • I understand the purposes for which the information collected about me can be used. • The revenue models of social media companies are familiar to me.
Self-Efficacy	Handling Data Collection (HDC) <ul style="list-style-type: none"> • I am able to influence the amount of information collected about me on social media. • I can influence the ways in which my personal information collected on social media is used commercially. • I actively influence how social media services obtain information about me.
	Defining Social Visibility (DSV) <ul style="list-style-type: none"> • I carefully define who gets to see my posts. • I often limit the visibility of my own profile on social media services. • I do not dare to put my personal information on social media.

	Category	Statements
Attitudes	Personal Data Privacy (PDP)	<ul style="list-style-type: none"> • I often worry about data privacy issues on social media. • I consider putting my personal information on social media a risk. • I am concerned that information about me and my posts can be used inappropriately. • Protecting the privacy of other users on social media is important to me. • I try to minimise the amount of information collected about me on social media. • I am willing to give up information about my private life in order to get free social media services.
	Privacy Trade- Offs (PTO)	<ul style="list-style-type: none"> • I find the collection of my user data on social media acceptable. • I think it is acceptable for information collected about me on social media to be resold. • Giving up information about my private life is a suitable compensation for free social media services. • Social media services have the right to sell personal information about me to fund their operations.

3.2 Data collection

Research data were collected from 163 mainly second-year pre-service teachers ($N_{\text{female}}=133$, 82%; $N_{\text{male}}=27$, 16%; $N_{\text{other}}=3$, 2%) who voluntarily participated from the University of Eastern Finland. The participants' ages ranged between 18 and 55 ($M_{\text{age}}=24.2$). Data were collected during a teacher training course named Societal Approaches in Science and History Education. The course falls under the umbrella of so-called multidisciplinary studies in Finnish teacher training covering the various subject areas taught in elementary schools (grades 1 to 6), including arts, music, history, mathematics, geography, Finnish and literature. Completing these multi-disciplinary studies is mandatory to teach students in grades 1 to 6. Data were collected via electronic questionnaires that contained 23 statements for six different areas (Table 1). Each participant was fully briefed on the aspects of their involvement, including ethical issues, methods of analysing the data and the purposes of the study. All participants gave consent. Participation to the study did not impact course completion or final grades. According to the guidelines of the Finnish National Board on Research Integrity (TENK), this study did not require prior ethics committee review, as all participants were adults, participation was voluntary, and informed consent was obtained within the questionnaire. Only limited demographic information, age, gender, and year of study were collected. The dataset was stored on an access restricted university server and handled in accordance with the University of Eastern Finland's data protection and GDPR guidelines (<https://www.uef.fi/en/data-protection>).

3.3 Data analysis

In order to assess the psychometric properties (RQ1) of the Data Agency instrument, the factor structure was evaluated by conducting a CFA using Jamovi 2.7.6 (Rosseel, 2012) software. The six-factor measurement model (see Appendix 1) was constructed to assess the instrument's structure. The model was estimated using the robust weighted least squares (WLSMV) method as it is suggested to be more suitable for Likert type variables and the data were moderate non normally distributed (Appendix 1. Table 2; Madria's skewness $p < 0001$, kurtosis $p = 0.059$). Furthermore, the model fit was assessed with robust fit indexes. The CFA fit of the model was assessed with the following goodness-of-fit indicators: the Tucker-Lewis index (TLI; Tucker & Lewis, 1973), the comparative fit index (CFI; Bentler, 1990), and the root mean square of the approximation (RMSEA; Steiger & Lind, 1980). A CFI or TLI larger than 0.95 was considered good, and 0.90–0.95 was considered mediocre (Hu & Bentler, 1999). An RMSEA (including 90% confidence intervals [90% CI]) smaller than 0.05 was considered excellent, and 0.06–0.08 was acceptable (Collier, 2020). To evaluate the internal consistency of the data agency factors, Cronbach's alpha (α) values were calculated for individual scale items. According to Nunnally and Berstein (1994), α approaching .70 or higher is acceptable. Furthermore, the skewness and kurtosis values were calculated to assess the normality of the constructed factors. In order to study pre-service teachers' perceptions of the data agency areas (RQ2), the descriptive statistics were reported, mean values and standard deviation. To further explore the differences among pre-service teachers (RQ3) k-means cluster analysis was conducted. K-means cluster analysis is typically conducted with different numbers of clusters to identify a solution that is the most meaningful for the purposes of the research (Jain, 2010). In this case, we tested two- and three-cluster solutions and ended up with a two-cluster model. Finally, to study the relations between data agency elements (RQ4), the Pearson correlation coefficients were calculated.

4. Results

The CFA measurement model for the six constructs (Appendix 1) achieved an acceptable, yet mediocre, fit after allowing some error terms to correlate. The fit indexes were $\chi^2 = 244$; $df = 215$; $P = 0.086$ and $\chi^2/df = 1.135$. As the chi-square test is sensitive to sample size (Collier, 2020), it was interpreted with caution. The robust fit indexes were TLI = 0.886, and CFI = 0.903, suggesting a mediocre fit (fit indexes near 0.9, Collier (2020); not reaching good fit 0.95, (Hu & Bentler 1999). The RMSEA = 0.057 (90% confidence interval: [0.039, 0.073]) with an RMSEA below 0.06 indicating a good fit (Hu & Bentler, 1999). The loadings of the items to the factors were mostly near 0.70, indicating that the model explained 50% of the variable's variance (Collier, 2020).

Cronbach's alpha value, as well as McDonald's ω , is between 0.67 and 0.77, indicating adequate internal consistency for each factor (Table 2). The skewness and kurtosis values settled between -1 and 1, showing that data aligns sufficiently with a normal distribution (Tähtinen et al., 2020).

Table 2. Cronbach alpha (α), skewness and kurtosis

Factors	α	ω	Skewness	Kurtosis
Recognising Influencing (RI)	.746	.784	-,348	,242
Understanding the Uses of Data (UUD)	.702	.725	,105	-,341
Privacy Trade-Offs (PTO)	.757	.754	,926	,650
Personal Data Privacy (PDP)	.678	.675	-,270	,229
Handling Data Collection (HDC)	.679	.701	,091	-,033
Defining Social Visibility (DSV)	.673	.664	-,601	-,068

The mean values were rather moderate, with a variance from 1.87 to 3.71, and the difference between the highest and lowest assessments was 1.86 (Table 3). Similarly, the standard deviations varied from 0.63 to 0.96, indicating relatively small differences between respondents.

Table 3. Means and standard deviations

	M (SD)
Recognising Influencing (RI)	3.71 (0.76)
Understanding the Uses of Data (UUD)	3.12 (0.85)
Privacy Trade-Offs (PTO)	1.87 (0.63)
Personal Data Privacy (PDP)	3.40 (0.73)
Handling Data Collection (HDC)	2.71 (0.78)
Defining Social Visibility (DSV)	3.57 (0.96)

Note. Likert-type scale (5=completely agree/4=agree/3=neither agree nor disagree/2=disagree/1=completely disagree).

The highest assessments (indicating areas of comparatively higher confidence) were for the factors Recognising Influencing (mean 3.71) and Defining Social Visibility (mean 3.57). However, even though these assessments scored the highest, the results showed that respondents were not very confident in these areas. Conversely, the lowest assessments were for the Privacy Trade-Offs (Mean 1.87) and Handling Data Collection (Mean 2.71). The Privacy Trade-Offs, with the lowest assessments, can be seen as a rather positive signal, suggesting that respondents value the privacy of their data. Altogether, it may be assumed that the mean values close to three within nearly all measured areas mean that the topics were not that familiar to the respondents, as shown by their cautious assessments.

The respondent profiles and differences between respondents can be seen more clearly within Figure 2 showing two clusters. The first cluster consists of 69 respondents and the second cluster 83 respondents. The two-cluster solution indicates that the assessments are almost at the same level for both clusters, i.e., the difference between respondents was small. Within the factors of Recognising Influencing, and Handling Data Collection, the differences are minimal. The biggest differences were in Defining Social Visibility sug-

gesting that pre-service teachers perceived differently their abilities to affect their appearance online, how their activities are shown for other users. Similarly, understanding the purposes of using the data collected from them causes more variation.

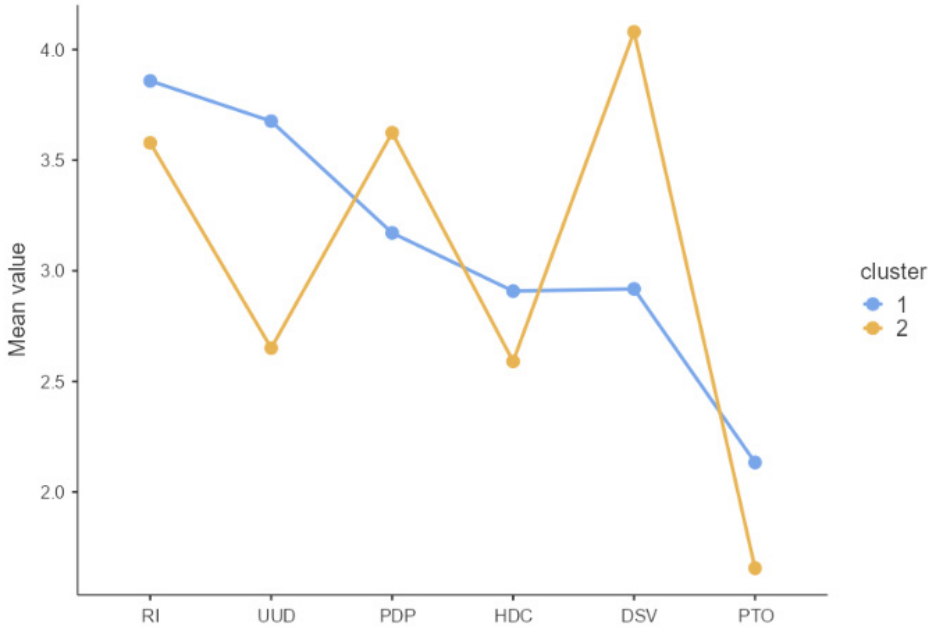


Figure 2. Data Agency Profiles.

Correlations between measured areas provide insights into the nature of data agency and how different areas are related to each other (Table 4). The correlations among the three main areas (understanding and awareness, self-efficacy and attitudes) are not as expected. The correlations between Recognising Influencing and Understanding the Uses of Data ($r=.45$) and between Personal Data Privacy and Privacy Trade-Offs ($r= -.45$) are as expected, and correlations are moderate. The two self-efficacy factors (Handling Data Collection and Defining Social Visibility) do not correlate, indicating that they are two separate areas. Interestingly, the strongest correlation ($r=.61$) is between Understanding the Uses of Data and Handling Data Collection. This suggests that these areas are perceived similarly. Also, a moderate correlation ($r = .36$) was found between Defining Social Visibility and Personal Data Privacy.

Table 4. Correlations between Data Agency areas

	Recognising Influencing	Understanding the Uses of Data	Privacy Trade-Offs	Personal Data Privacy	Handling Data Collection
Understanding the Uses of Data	.45**				
Privacy Trade-Offs	.06	.21*			
Personal Data Privacy	.24*	-.04	-.45**		
Handling Data Collection	.25*	.61**	-.10	.29*	
Defining Social Visibility	.10	-.10	-.14*	.36**	.06

** . Correlation is significant at the 0.01 level, * significant at 0.05 level (2-tailed).

5. Discussion

This paper aimed to introduce a questionnaire for measuring pre-service teachers' data agency and provide insights on their self-assessments from this perspective. The questionnaire was developed based on previous studies and definitions of data agency (Tedre et al., 2020; Vartiainen et al., 2021; Kahila et al., 2024). Three main areas (understanding, self-efficacy and attitudes) were identified and operationalised via six factors (23 statements). The CFA achieved a mediocre fit only after allowing some error terms to correlate. This was especially the case in the Personal Data Privacy and Privacy Trade-Offs factors, which may indicate that the model as presented is not fully capturing these constructs and that some new statements/items are needed. Also, certain items had quite low loadings (Appendix 1), indicating low item-level reliability. Interestingly, those items were quite general level statements such as: 'In social media, I often notice content that I recognise as being intentionally targeted at me' in the factor Recognising Influence and 'I do not dare to put my personal information on social media' in the factor Defining Social Visibility. The instrument could benefit if those items were as specific as others in those factors.

According to Tedre et al. (2020), people are often unaware of how and for what purposes their data can be used. Carmi et al. (2020) and Wenger (2023) note that it is highly important for people to identify different attempts to influence their activities and choices. The results of this article suggest that pre-service teachers were rather consistent in their responses, with small variations. Respondents were somewhat confident in their ability to recognise different influencing attempts. Still, the assessment of how well they understood the overall mechanisms of datafication was lower. This is understandable, as pre-service teachers are more confident with recognising influencing attempts, such as content generated especially for them. These attempts might be part of respondents' everyday experiences, recognising that different online contents had already aligned with their previous searches or articles they had read. Again, the invisible part related to data, its collection, analysis and profiling, remained invisible (Tedre et al., 2020), demanding more conscious research and familiarisation with the topic. From the perspective of computing education, these results point toward the need for targeted teacher-education interventions that explicitly develop pre-service teachers' data agency alongside core computing concepts. Without sufficient data agency, future teachers may struggle to move beyond surface-level

messages about digital safety to engage learners with the computational processes shaping their online lives.

When considering self-efficacy as part of data agency, previous studies emphasised that young people seem rather skilled in shaping online visibility and determining who can access their content (Davis & James, 2013; Hargittai & Marwick, 2016). Again, the challenge has been achieving self-efficacy, and actual skills, regarding commercial tracking and collection of data (Keen, 2020; Lapenta & Jørgensen, 2015). The result aligns with previous findings, with the difference being that self-efficacy affects social presence more than commercial data collection. The factor measuring the ability to affect social visibility was the second-highest data agency area, while the ability to affect data collection for commercial purposes was lower. Aligning with Lapenta and Jørgensen (2015) and Lievens et al. (2019), it appears that, for respondents, the results of taking care of their social presence online are more important and more relevant on an ongoing basis than the more invisible effect resulting from personalised content.

The results indicate that respondents valued their data. Previous studies suggested that young people often perceive the use of their data for commercial purposes as acceptable (Lapenta & Jørgensen, 2015; Selwyn & Pangrazio, 2018), and it has even been seen as an obligation of the platform providers (Selwyn & Pangrazio, 2018). Compared to these results, the findings seem rather contradictory. The value of personal data was perceived as high, and it seemed that pre-service teachers did not view their data as a suitable exchange for free applications. However, the role of personal data privacy was perceived more moderately, suggesting that the challenges and risks of publishing personal data online did not create much concern. However, people who are unaware of the system may not be aware of its possible risks.

The correlations between data agency areas showed that some areas were assessed similarly, suggesting greater similarities in the lives of pre-service teachers. Correlations also indicated that understanding the uses of data was positively related to recognizing, influencing, and handling the data collected. This highlighted the importance of understanding the areas that are not readily visible to everyday social media users. Consistent with Lievens et al. (2019), the focus and target of acknowledged decisions (the data agency) seemed to center on current issues directly visible to users. Still, this area needs more research to better understand how pre-service teachers perceive data agency areas and, especially, how the areas affect one another. A more diverse set of methods would be appropriate for future efforts.

Lupton and Williamson (2017) suggested that young people living in technology-rich environments gain a good starting point for developing their data agency during their childhood. The mean age of the respondents in this study was 24, indicating most were born at the beginning of 2000 and have lived their lives during web 2.0, social software and eventually datafication. However, assessments were rather cautious, mainly at the middle of the scale, suggesting that the areas measured are not overly familiar to pre-service teachers, as responses did not indicate a strong stance. Still, data agency and datafication are increasingly important roles (Valtonen et al., 2019). The support young people receive from their

homes regarding the data agency varies greatly (Livingstone & Yoo, 2018; Stoilova et al., 2020). Different levels of support may lead to inequalities in coping with the challenges of the digitalised society. This is true even in situations where teachers do include the challenges and risks related to online safety in their work, usually focusing on e-safety, but ignore concepts like tracking, profiling, targeted advertisement, recommender systems or other everyday techniques that drive the data economy (Stoilova et al., 2019; Valtonen et al., 2019). The results of this study indicate that future support for young people and pre-service teachers will be appropriate to facilitate their improved understanding of the invisible areas of datafication and the mechanisms, effects and possibilities of commercial data collection. This is a key area for developing a data agency, providing young people with the means to better contribute to their own lives and ensuring they remain independent of algorithms that attempt to make decisions on their behalf.

The role of data agency needs to be more prominently highlighted within teacher training. The topic should go beyond technical skills to address the often-invisible aspects of datafication and commercial data collection. Teacher education must help pre-service teachers understand how data flows, what mechanisms drive algorithmic decisions, and the societal implications of these processes, not only for their professional practice but also for their personal lives. Achieving this requires a holistic approach that combines courses in the social sciences and critical literacy with modules focused on digital technologies. Tools such as *Somekone* (<https://somekone.gen-ai.fi/start>) can provide valuable support for students and pre-service teachers to understand the mechanisms of social software, data collection, and recommendation systems (see Pope et al., 2025). In the future, there is a need to develop more of this kind of technology and related pedagogies. The aim is to ensure that future teachers can provide their students with a nuanced understanding of digitalization, its possibilities, and its risks for their personal agency, equipping them with the ability to make informed choices and decisions in environments increasingly shaped by algorithmic systems and data-driven platforms.

This study provides interesting perspectives on the data agency of pre-service teachers and on how to measure such agency. In the future, the questionnaire introduced will require further development and testing, especially regarding the two statements with low loadings. Furthermore, the Personal Data Privacy factor could be further developed to better highlight possible apathy (see Selwyn & Pangrazio, 2018) towards the potential for affecting commercial data collection.

In the future, it will be important to continue researching the data agency, especially in the context of teacher education. In addition, further research and understanding are needed about the nature of pre-service teachers' data agency and ways to support data agency during teacher training. Altogether, the findings of this study should be interpreted with caution, as it relied on a relatively small, homogeneous sample of Finnish pre-service teachers. Also, the instrument demonstrated only a mediocre model fit in confirmatory factor analysis, suggesting that some constructs may not be fully captured and may require further refinement to improve validity. Nevertheless, this article represents an important step toward understanding pre-service teachers' abilities to navigate a datafied society.

6. Conclusion

This study provided insights into the nature of data agency among pre-service teachers and the importance of further studying this area. The developed questionnaire serves as a starting point for this research area but still needs further development. Findings suggest that while pre-service teachers show a reasonable level of confidence across most data agency areas, they still face challenges in understanding the underlying datafication processes and associated commercial data practices. The invisible aspects of datafication need to be highlighted, and pre-service teachers' understanding of these mechanisms should be supported. These results show the need for targeted educational interventions and improvements in teacher training programs to foster pre-service teachers' stronger data agency and their capability to consider these themes within their future work in schools. From a computing education perspective, the data agency provides a framework that connects ethical, societal, and pedagogical concerns with an understanding of the technical mechanisms that drive datafication.

Acknowledgments

The authors would like to extend their gratitude to the January Collective for its continued support.

Generative AI Disclosure: The authors used the generative AI tool Copilot only for linguistic refinements. The scientific content, analysis and interpretations were done solely by the authors.

Author Contributions

Teemu Valtonen: Conceptualization, methodology and writing the original draft.

Henriikka Vartiainen: Conceptualization and writing the original draft.

Eliisa Vähä: Data collection and writing the original draft.

Jari Kukkonen: Methodology and formal analysis.

Erkko Sointu: Methodology and formal analysis.

Juho Kahila: Data collection and writing the original draft.

Matti Tedre: Conceptualization, writing the original draft and writing review & editing.

All authors have read and approved the final manuscript.

References

- Acquisti, A., Brandimarte, L. & Loewenstein, G. (2015). Privacy and human behavior in the age of information. *Science*, 347(6221), 509–514. DOI: 10.1126/science.aaa1465.
- Baack, S. (2015). Datafication and empowerment: How the open data movement re-articulates notions of democracy, participation, and journalism. *Big Data & Society*, 2(2). DOI: 10.1177/2053951715594634.
- Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, 37(2), 122.
- Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, 52, 1–26.
- Bawden, D. (2001). Information and digital literacies: A review of concepts. *Journal of Documentation*, 57(2), 218–259. DOI: 10.1108/EUM0000000007083.
- Bentler, P. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107(2).
- Berland, M. and Garcia, A. (2024). *The Left Hand of Data: designing Education Data for Justice*. The MIT Press, Cambridge, MA, USA.
- Boyd, d. (2015). *It's Complicated: The Social Lives of Networked Teens*. New Haven: Yale University Press.
- Boyd, d. & Ellison, N. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13(1), 210–230.
- Boyd, d. & Marwick, A. (2011). Social privacy in networked publics: Teens' attitudes, practices, and strategies. In *A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society*, Oxford Internet Institute, 22 September 2011, pp. 1–29. Available at: papers.ssrn.com/sol3/papers.cfm?abstract_id=1925128 (accessed XX Month XXXX).
- Bowker, G. C. & Star, S. L. (2000). *Sorting Things Out: Classification and Its Consequences*. Cambridge: The MIT Press.
- Bowler, L., Acker, A., Jeng, W. & Chi, Y. (2017). 'It lives all around us': Aspects of data literacy in teens' lives. *Proceedings of the Association for Information Science and Technology*, 54(1), 27–35.
- Carmi, E., Yates, S. J., Lockley, E. & Pawluczuk, A. (2020). Data citizenship: Re-thinking data literacy in the age of disinformation, misinformation, and malinformation. *Internet Policy Review*, 9(2).
- Castells, M. (1996). *The Rise of the Network Society (The Information Age: Economy, Society and Culture, Volume 1)*. Blackwell Publishing.
- Collier, J. (2020). *Applied structural equation modeling using AMOS: Basic to advanced techniques*. Routledge.
- Couldry, N. & Powell, A. (2014). Big data from the bottom up. *Big Data & Society*, 1(2).
- Culbertson, J. (1986). Whither computer literacy? In J. A. Culbertson & L. L. Cunningham (Eds.), *Microcomputers and Education (Vol. 1, pp. 109–131)*. The University of Chicago Press.
- Cui, Y., Chen, F., Lutsyk, A., Leighton, J. P. & Cutumisu, M. (2023). Data literacy assessments: A systematic literature review. *Assessment in Education: Principles, Policy & Practice*, 30(1), 76–96.
- Davis, K. & James, C. (2013). Tweens' conceptions of privacy online: Implications for educators. *Learning, Media and Technology*, 38(1), 4–25.
- diSessa, A. (2000). *Changing Minds: Computers, Learning, and Literacy*. The MIT Press.

- Emirbayer, M. & Mische, A. (1998). What is agency? *The American Journal of Sociology*, 103(4), 962–1023.
- Eteläpelto, A., Vähäsantanen, K., Hökkä, P. & Paloniemi, S. (2013). What is agency? Conceptualizing professional agency at work. *Educational Research Review*, 10, 45–65. <https://doi.org/10.1016/j.edurev.2013.05.001>
- Finnish National Board of Education. (2014). *Perusopetuksen opetussuunnitelman perusteet 2014* [Finnish national core curriculum for basic education]. Retrieved from https://www.oph.fi/sites/default/files/documents/perusopetuksen_opetussuunnitelman_perusteet_2014.pdf
- Harel, D. & Feldman, Y. (2004). *Algorithmics: The Spirit of Computing* (3rd ed.). Addison-Wesley.
- Hargittai, E. & Marwick, A. (2016). ‘What can I really do?’ Explaining the privacy paradox with online apathy. *International Journal of Communication*, 10, 3737–3757.
- Hendricks, V. F. & Vestergaard, M. (2018). *Reality Lost: Markets of Attention, Misinformation and Manipulation*. Springer.
- Hintz, A., Dencik, L. & Wahl-Jorgensen, K. (2018). *Digital citizenship in a datafied society*. John Wiley & Sons.
- Hu, L. & Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, 6(1).
- Höper, L. & Schulte, C. (2023). The data awareness framework as part of data literacies in K-12 education. *Information and Learning Sciences*. <https://doi.org/10.1108/ILS-06-2023-0075>.
- Höper, L. and Schulte, C. (2024). Empowering students for the data-driven world: A qualitative study of the relevance of learning about data-driven technologies. *Informatics in Education*, 23(3), 593–624.
- Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, 31(8), 651–666.
- Keen, C. (2020). Apathy, convenience or irrelevance? Identifying conceptual barriers to safeguarding children’s data privacy. *New Media & Society*, 24, 50–69. <https://doi.org/10.1177/1461444820960068>
- Kennedy, H., Poell, T. & van Dijck, J. (2015). Data and agency. *Big Data & Society*, 2(2). <https://doi.org/10.1177/2053951715621569>
- Kitchin, R. (2012). *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. SAGE.
- Kramer, A. D. I., Guillory, J. E. & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences of the United States of America*, 111(24), 8788–8790.
- Kubin, E. & von Sikorski, C. (2021). The role of (social) media in political polarization: a systematic review. *Annals of the International Communication Association*, 45(3), 188–206.
- Lapenta, G. H. & Jørgensen, R. F. (2015). Youth, privacy and online media: Framing the right to privacy in public policy-making. *First Monday*, 20(3). <https://doi.org/10.5210/fm.v20i3.5568>
- Lehtiniemi, T. & Haapoja, J. (2020). Data agency at stake: MyData activism and alternative frames of equal participation. *New Media & Society*, 22(1), 87–104. <https://doi.org/10.1177/1461444819861955>
- Lievens, E., Livingstone, S., McLaughlin, S., O’Neill, B. & Verdoodt, V. (2019). Children’s rights

- and digital technologies. In U. KilKelly & T. Liefwaard (Eds.), *International Human Rights of Children* (pp. 487–513). Springer.
- Livingstone, S. & Yoo, D. (2018). More clarity brings more confusion: debating what the European General Data Protection Regulation means for children's right to protection from harm online. *New Media & Society*, 20(5), 1622–1638. <https://doi.org/10.1177/1461444817745012>
- Lupton, D. & Williamson, B. (2017). The datafied child: The dataveillance of children and implications for their rights. *New Media and Society*, 19(5), 780–794. <https://doi.org/10.1177/1461444816686328>
- Marwick, A. E. & boyd, d. (2014). Networked privacy: how teenagers negotiate context in social media. *New Media and Society*, 16(7), 1051–1067.
- Miao, F. (2022). K–12 AI Curricula: A Mapping of Government-Endorsed AI Curricula. UNESCO, Paris, France.
- Miao, F. and Shiohira, K. (2024). AI Competency Framework for Students. UNESCO, Paris, France.
- NRC Committee on Information Technology Literacy. (1999). *Being Fluent with Information Technology*. National Academy Press.
- Nunnally, J. & Bernstein, I. (1994). *The Assessment of Reliability*. *Psychometric Theory*, 3.
- Pangrazio, L. & Selwyn, N. (2019). 'Personal data literacies': a critical literacies approach to enhancing understandings of personal digital data. *New Media and Society*, 21(2), 419–437.
- Papert, S. (1980). *Mindstorms: Children, Computers, and Powerful Ideas*. Basic Books.
- Pelletier, M. J., Krallman, A., Adams, F. G. & Hancock, T. (2020). One size doesn't fit all: A uses and gratifications analysis of social media platforms. *Journal of Research in Interactive Marketing*, 14(2), 269–284.
- Pope, N., Kahila, J., Vartiainen, H., Saqr, M., López-Pernas, S., Roos, T., ... & Tedre, M. (2025, April). An XAI Social Media Platform for Teaching K-12 Students AI-Driven Profiling, Clustering, and Engagement-Based Recommending. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 39, No. 28, pp. 29203–29211).
- Raynes-Goldie, K. (2010). Aliases, creeping, and wall cleaning: Understanding privacy in the age of Facebook. *First Monday*, 15(1), 2775.
- Ridsdale, C., Rothwell, J., Smit, M., Ali-Hassan, H., Bliemel, M., Irvine, D., Wuetherick, B. & Wuetherick, B. (2015). *Strategies and best practices for data literacy education: Knowledge synthesis report*. Dalhousie University.
- Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(2), 1–36. <https://doi.org/10.18637/jss.v048.i02>
- Selwyn, N. & Pangrazio, L. (2018). Doing data differently? Developing personal data tactics and strategies amongst young mobile media users. *Big Data and Society*, 5(1). <https://doi.org/10.1177/2053951718765021>
- Sheil, B. (1980). Teaching procedural literacy. In *Proceedings of the ACM 1980 Annual Conference* (pp. 125–126). ACM.
- Smith, H. J., Dinev, T. & Xu, H. (2011). Information privacy research: An interdisciplinary review. *MIS Quarterly*, 35(4), 989–1015.
- Steiger, J. H. (1980). Statistically based tests for the number of common factors. In Paper presented at the Annual Meeting of the Psychometric Society, Iowa City, 1980.
- Stoilova, M., Livingstone, S. & Nandagiri, R. (2019). Children's data and privacy online: growing up in a digital age: research findings.

- Stoilova, M., Livingstone, S. & Nandagiri, R. (2020). Digital by default: Children's capacity to understand and manage online data and privacy. *Media and Communication*, 8(4), 197–207. <https://doi.org/10.17645/mac.v8i4.3407>
- Swan, M. (2012). Sensor mania! the Internet of Things, wearable computing, objective metrics, and the quantified self 2.0. *Journal of Sensor and Actuator Networks*, 1(3), 217–253.
- Tähtinen, J., Laakkonen, E. & Broberg, M. (2020). *Tilastollisen aineiston käsittely ja tulkinnan perusteita* [Analysis and interpreting of statistical data]. Publications University of Turku, Faculty of Education C, 22 (2nd edition).
- Tedre, M. & Denning, P. (2016, November). The long quest for computational thinking. In *Proceedings of the 16th Koli Calling international conference on computing education research* (pp. 120–129).
- Tedre, M., Vartiainen, H., Kahila, J., Toivonen, T., Jormanainen, I., and Valtonen, T. (2020). Machine learning introduces new perspectives to data agency in K–12 computing education. In *2020 IEEE Frontiers in Education Conference (FIE)*, 1–8.
- Tucker, L. & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38(1).
- Tufekci, Z. (2008). Can you see me now? Audience and disclosure regulation in online social network sites. *Bulletin of Science, Technology & Society*, 28(1), 20–36.
- Valtonen, T., Tedre, M., Mäkitalo, K. & Vartiainen, H. (2019). Media Literacy Education in the Age of Machine Learning. *Journal of Media Literacy Education*, 11(2). <https://doi.org/10.23860/jmle-2019-11-2-2>
- Vartiainen, H., Pellas, L., Kahila, J., Valtonen, T., and Tedre, M. (2024). Pre-service teachers' insights on data agency. *New Media & Society*, 26(4):1871–1890.
- Vartiainen, H., Toivonen, T., Jormanainen, I., Kahila, J., Tedre, M., and Valtonen, T. (2021). Machine learning for middle schoolers: Learning through data-driven design. *International Journal of Child-Computer Interaction*, 29:100281.
- van Dijck, J. (2013). 'You have one identity': performing the self on Facebook and LinkedIn. *Media, Culture and Society*, 35(2), 199–215
- Vee, A. (2010). *Proceduracy: Computer Code Writing in the Continuum of Literacy*. (PhD thesis). University of Wisconsin at Madison.
- Wenger, E. J. (2023). *Reclaiming Data Agency in the Age of Ubiquitous Machine Learning*. (PhD Thesis). The University of Chicago.
- Wertsch, J. V. (2007). Mediation. In H. Daniels, M. Cole & J. V. Wertsch (Eds.), *The Cambridge Companion to Vygotsky* (pp. 178–192). Cambridge University Press.
- Wing, J. M. (2006). Computational thinking. *Communications of the ACM*, 49(3), 33–35.
- Zuboff, S. (1988). *In the age of the smart machine: The future of work and power*. Basic Books, Inc.
- Zuboff, S. (2015). Big other: surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, 30(1), 75–89.
- Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. Profile Books.

Appendix 1.

Figure 3. The six-factor measurement model.

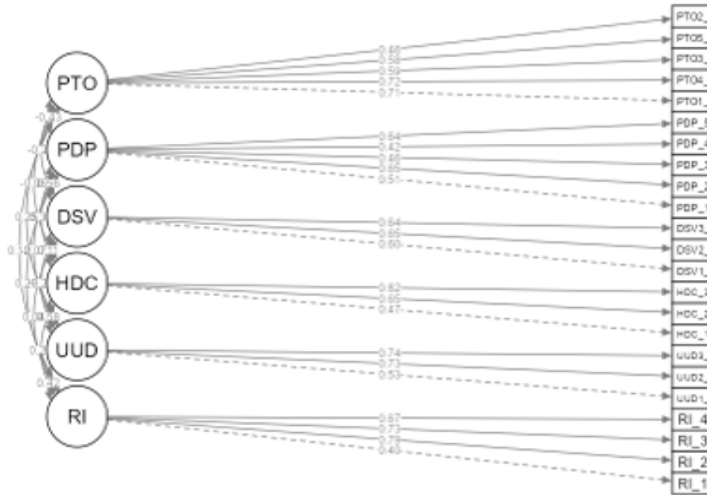


Table 5. Item-level descriptive statistics with skewness and kurtosis values

	Mean	SD	Skewness		Kurtosis	
			Skewness	SE	Kurtosis	SE
RI_1	4.52	0.707	-2.09611	0.191	6.8255	0.379
RI_2	3.20	1.121	0.00348	0.190	-0.8931	0.378
RI_3	3.49	1.047	-0.41084	0.191	-0.2872	0.379
RI_4	3.63	1.031	-0.63405	0.190	-0.1939	0.378
HDC_1	2.93	0.898	0.14786	0.191	-0.0799	0.380
HDC_2	2.61	1.068	0.28518	0.190	-0.4470	0.378
HDC_3	2.60	1.018	0.42719	0.191	-0.1644	0.379
PDP_1	3.06	1.115	-0.01399	0.190	-0.6430	0.378
PDP_2	3.68	1.098	-0.61178	0.190	-0.3730	0.378
PDP_3	3.21	1.119	-0.12685	0.190	-0.7026	0.378
PDP_4	3.97	0.929	-0.88058	0.191	0.9340	0.379
PDP_5	3.11	1.100	-0.02421	0.190	-0.6152	0.378
PTO_1	2.12	1.036	0.61068	0.191	-0.3579	0.379
PTO_2	2.19	0.912	0.40909	0.191	-0.5843	0.380
PTO_3	1.55	0.818	1.40799	0.190	1.1962	0.378
PTO_4	1.78	0.819	0.62945	0.191	-0.6409	0.380
PTO_5	1.63	0.833	1.17873	0.191	0.5758	0.379
DSV_1	3.71	1.232	-0.68399	0.190	-0.5261	0.378
DSV_2	3.88	1.173	-0.85582	0.191	-0.2398	0.379
DSV_3	3.13	1.276	0.00802	0.191	-1.1396	0.379
UUD_1	3.61	0.871	-0.33494	0.190	-0.2628	0.378
UUD_2	3.09	1.111	0.01997	0.191	-0.6819	0.379
UUD_3	2.66	1.188	0.24725	0.190	-0.8753	0.378