Report on Quantitative Analysis of Survey Data

Work package n°2 - Training Plan: Learning Hurdles

Leading organisation - University of Education, Freiburg

Deborah Krzyzowski, with Bernd Remmele, Helena Strauß, and Zlatko Valentić

Co-funded by the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor EACEA can be held responsible for them.

Content

I. Introduction

For development of user-based learning tools, a crucial step is identification of potential *learning hurdles* to students. To evaluate the current state of students' awareness of privacy issues online, and their knowledge of concrete measures to protect their privacy, the partner organisations designed an online survey. Collected answers will be discussed in this report and may serve as indicators of potential obstructions to learning success.

The objectives of *DataPro* encompass education on data protection for students, and supporting teachers through providing learning tools developed in *MeCyS*. During the project duration, these tools will undergo consistent adaption and development. The results of the data analysis presented in this report may reveal the students' needs for, and perspective on data protection to be met for an user-based experience.

First, the research framework will be presented to next explain derived hypotheses in the research design. The methods of hypothesis testing will be laid out and the results visualized, and later connected with relevant literature. The *DataPro Curriculum* laid the foundations for the survey questions, as sketched out in following chapter.

II. Research Framework

The *DataPro Curriuclum* identifies three core areas for the learning tools on data protection. These are a. Data Protection, b. Privacy as a Right, and c. Data as a Good. They should reflect 'the requirements of an active, autonomous, and responsible member of a democratic, data/information-driven market society' (Remmele and Valentic, 2024, p. 1). These core areas of focus include thematic subcategories – like data privacy risks, privacy policies, or the trading of data. The subcategories intertwine with the education goals of *DataPro*, which are teaching students about data protection measures, and further, on their individual privacy rights as elementary to healthy democracies. The relations between such subcategories are schematically exemplified in the following chart (Figure 1).

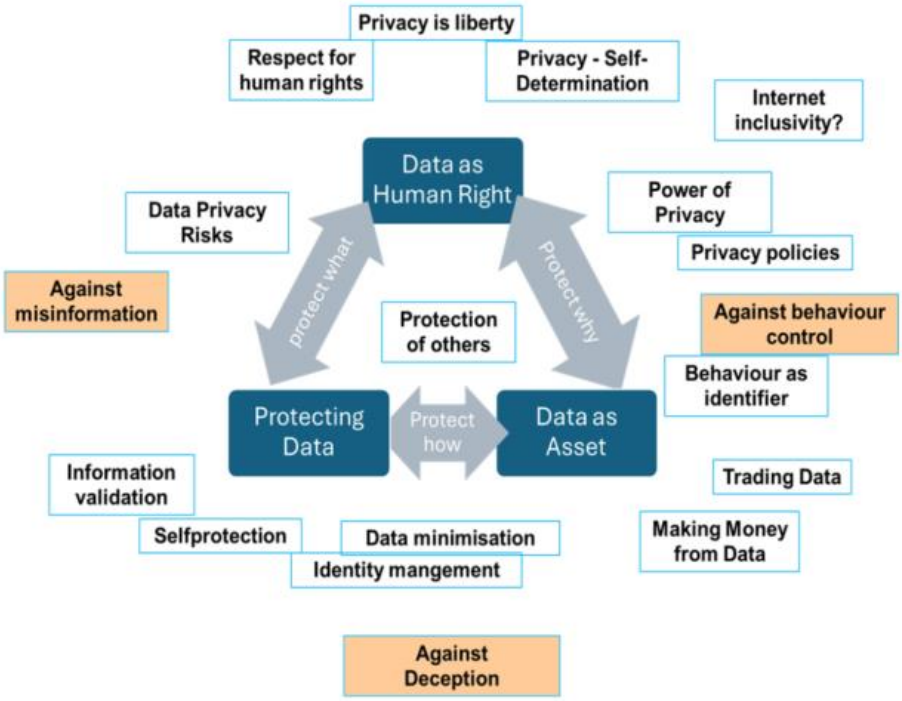


Figure 1. **Competence Matrix on Data Protection.**

III. Research Design

Based on the *Competence Matrix*, the survey questions estimated students' knowledge of, and attitude towards data protection and online privacy (see Appendix A for detailed thematic organisation of the survey questions). Different answer-modes of either open-text format, of level of agreement based on a 5-point Likert scale, or as multiple-choice selection measured their experience in the survey. the Questions for example investigated the participants' knowledge on concrete measures for data protection, like what a safe password looks like or what browser-cookies do (question 10 and 17). Additionally, the participants demographical data on age, country of residence, and gender was collected on a voluntary basis, and their internet usage habits queried.

The three core topics of the *Competence Matrix* allows for clustering of the survey questions as investigating the student's a. practical knowledge of data protection; b. their awareness of their privacy rights; and c. their awareness of the monetary and societal value of their data. The answers could then be translated into variables measuring for example, students' trust in online environments, or their data protection skills. In the following table, the survey questions are sorted in named clusters. Such categorization allows for their later operationalization as variables in the quantitative data analysis (see Table 1).

Relationship between these variables may be inferred from significant differences between two population means, which will be explained in the upcoming methods section. Together with the participants' demographic data, the survey data analysis will give insight into the students' general tendencies towards topics within the three clusters A-C (Table 1). A descriptive overview of the survey results will be given next, and derived hypotheses presented.

A. *Protecting Data* - Data Protection Awareness
    a. Data protection skills regarding
        i. Phishing
        ii. Password security
        iii. Managing accounts on mobile devices
    b. Knowledge on cyber security

A. *Data as Human Right* - Privacy Rights Awareness
    c. Knowledge of legal rights under e.g., GDPR
    d. Trust in online environments
    e. Perception of public-private dichotomies

B. *Data as Asset* - Data Value Awareness
    f. Knowledge on data processing methods for profit
    g. Perception of value of personal data

Table 1. **Operationalisation of Competencies.**

a. Data Collection

The sample data consists of all completed surveys. The survey itself was designed with the open-source and anonymous online survey tool LimeSurvey (Nagel, 2024). Contacted teachers shared the link or QR-Code to the survey with their students in class. Previously, the teachers at partaking schools had been informed about the objectives of the survey and its design in online-meetings. Answers were

collected in several rounds, first starting in early June 2024, and running last in early August 2024. Any answers entered after 12th of August 2024 were not included in the first descriptive overview.

Participation in the survey was encouraged by the teachers yet happened on a voluntary basis. No rewards were given out upon completion of the survey. Answers to every question of the survey were obtained in a no-pressure environment, and questions could be skipped without being answered. Disclosure of demographic data was also made on a voluntary basis, with the options to skip or 'prefer not to say'.

b. Descriptive Overview: Knowledge on Data Protection

First sighting of the data revealed that a majority of 69% of the participants report daily use of the internet without restrictions (Figure 2). Over 80% of the participants have profiles on four or more social media accounts, with Instagram and TikTok being the most popular platform.

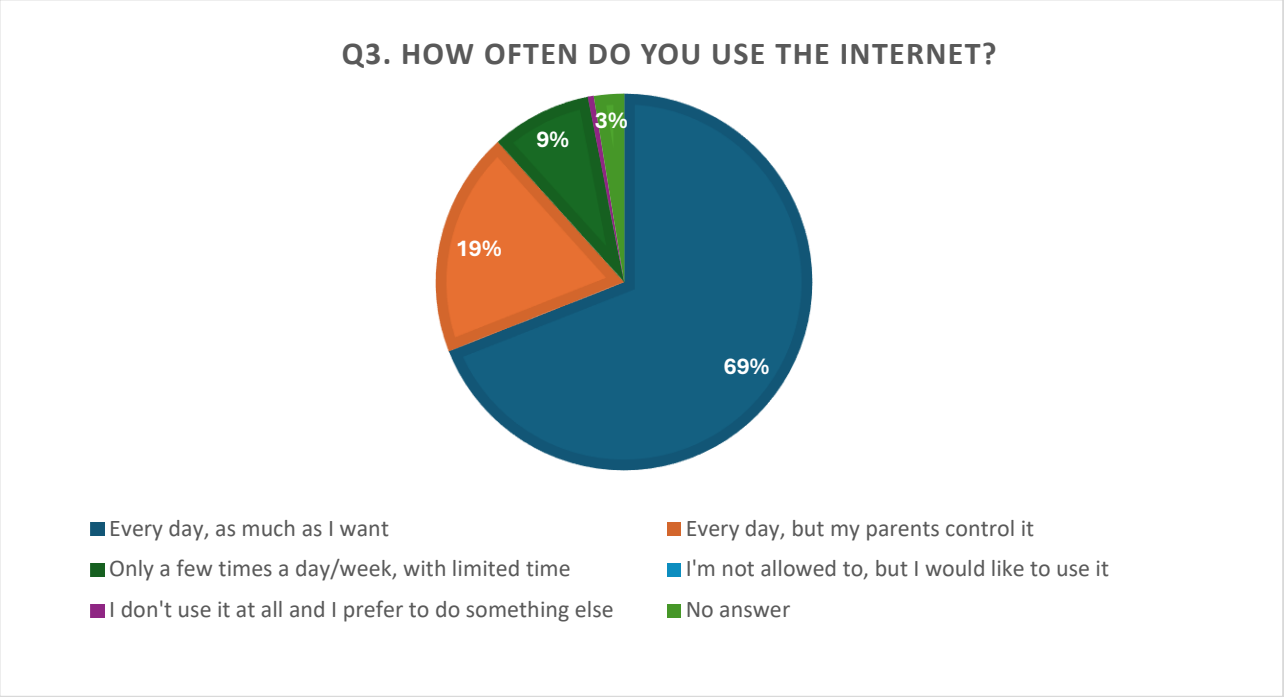


Figure 2. Overview of Answers to Question 3.

Regarding the participants' reported cyber security awareness, knowledge on the meaning of cyber security threats varied. About 49% of all answers reported a definite 'yes' when asked if they knew what was meant by 'identity theft'. In contrast, such knowledge sunk to 38% when asked about 'phishing'. Awareness of the own vulnerability to cyber attacks was positive for about a half (47%) of participants who agreed or strongly agreed that their own social media accounts may be targeted (Figure 3). It is hence of interest to investigate the factors which may influence lack of, or a high awareness of cyber security risks. As the reported age of the participants is mostly clustered at 13 to 15 years, Hypothesis A. investigates whether the participants' age may be a factor for change in cyber

security awareness, and trust in online environments. At a significance level above 0.05, H_{0A}: 'Mean level of cyber security awareness does not change with participant's age' may be accepted.

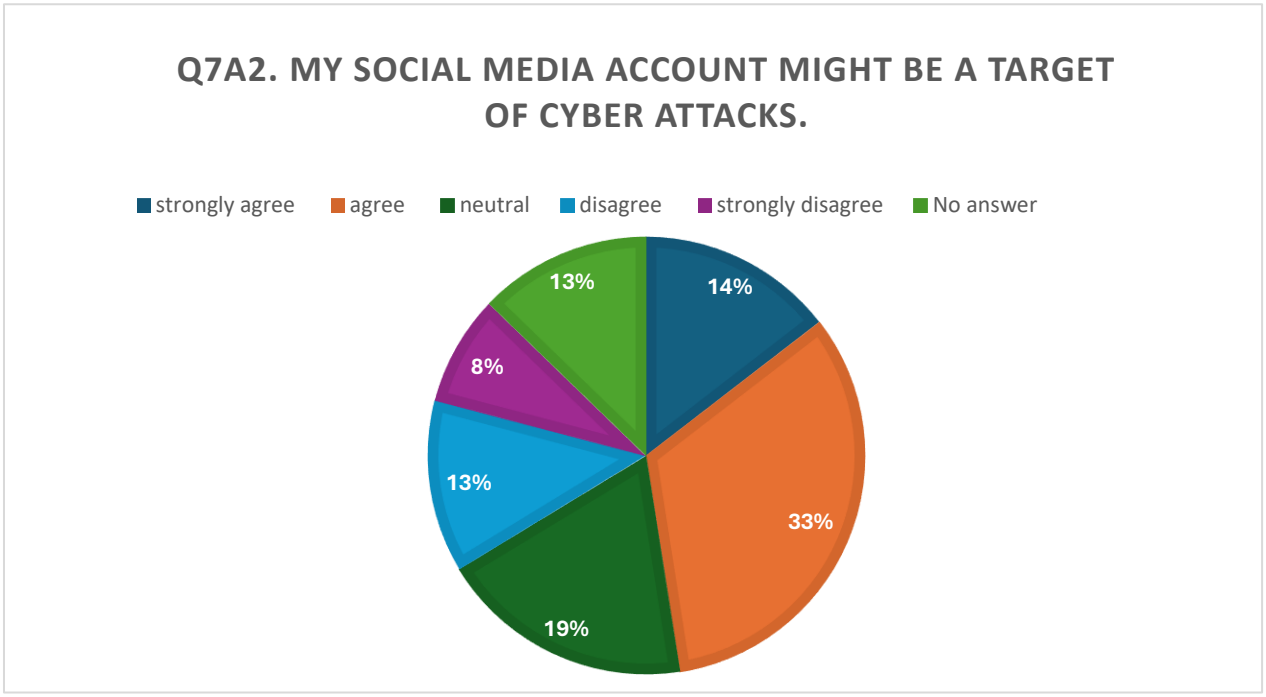


***Figure 3*. Overview of Answers to Question 7.A2.**

c. Descriptive Overview: Trust in Online Environments, and Attitude towards Privacy

Concerning participants' trust in online environments, the majority disagree or strongly disagree that

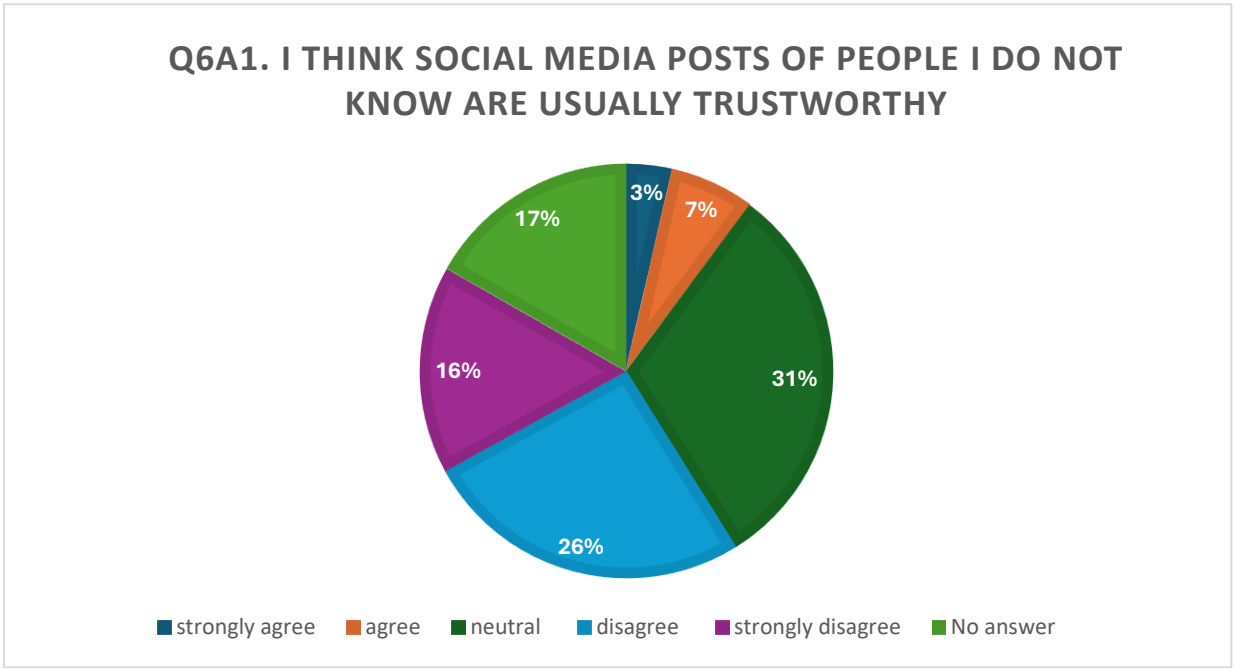


Figure 4. Overview of Answers to Question 6. A1.

social media posts of people they do not know are trustworthy (Figure 4). Previous studies found correlations between online behaviour, as well as between gender and adolescents' awareness of online risks (Savoia *et al.*, 2021, p. 12). Beyond the participants' age, change in mean levels of trust in online environments and their attitude towards online privacy might be observed with a change in indicated gender. In the case of significant change in mean levele of agreements to statements of trust and privacy H_{0B}: 'Mean level of trust in online environments, and privacy attitudes do not change with the participant's gender' may be rejected.

d. Descriptive Overview: Online Usage Habits, and Users' Preferences

Previously, online behaviour was briefly mentioned as an explaining factor for participants' level of awareness of cyber security threats. Next to demographic variables of a. age, and b. gender, the effect of c. online behaviour on a change in these variables is to be assessed. The survey queried online behaviour in Question 3: 'How often do you use the Internet?' (see Figure 2). The answers were operationalised as internet usage habits. Participants were then grouped into three groups: everyday users, everyday users under parental oversight, or rare users. Data protection awareness and attitudes to online privacy were queried in Question 19 with statements like 'I am not worried about my online data; I have nothing to hide'. If a significant change of $p < .05$ in mean levels of agreement to such statement between respective groups of habits is observed, usage habits may influence the general attitude towards online privacy or vice versa. Consequently, the alternative hypothesis H_{1C}: 'Mean levels of privacy attitudes do change with online usage habits of the participant' may be accepted.

To determine whether other factors may have an influence on participants' attitude towards privacy issues and data protection, their preference for ease of access was investigated. Preferring ease of access was theorized as alternative explanation to indifference towards online privacy. Students may prioritize quick log-on options or manageability of passwords while being aware of privacy issues. To explore this alternative explanation, agreement levels to the 'nothing to hide' argument in question 19 served as indicator of lower awareness of privacy issues. Online habits like the use of password managers and linking social media accounts to other online services was evaluated as a preference for ease of access (see questions 11 and 12, respectively). H_{1D}: 'Mean levels of privacy attitudes change with the participant's online preferences' tests this potential relation and is to be accepted if $p< .05$.

Summarized, the observed change in means between grouped populations may give insight into the relationship between demographic data and the three areas of data protection competencies (see Table 1). Next to age and gender, non-demographic data like online usage habits, and online preferences may help to find factors which affect data protection awareness, or the attitude towards

online privacy. Before presenting the methods of analysis, Table 2. offers an overview of the research's hypotheses.

Competencies A – Data Protection Awareness

Hypotheses A0: Mean level of cyber security awareness does not change with participant's age.

Hypotheses A1: Mean level of cyber security awareness changes with participant's age.

Competencies B - Privacy Rights Awareness

Hypotheses B0: Mean level of trust in online environments, and privacy attitudes do not change with the participant's gender.

Hypotheses B1: Mean level of trust in online environments, and privacy attitudes change with the participant's gender.

Hypotheses C0: Mean levels of privacy attitudes do not change with online usage habits of the participant.

Hypotheses C1: Mean levels of privacy attitudes change with online usage habits of the participant.

Competencies C - Data Value Awareness

Hypotheses D0: Mean levels of privacy attitudes do not change with the participant's online preferences.

Hypotheses D1: Mean levels of privacy attitudes change with the participant's online preferences.

Table 2. **Research Hypotheses and Corresponding Competencies.**

IV. Methods

To compare the change in mean level of agreement to statements, the Likert scale of 5 levels of agreement from 'strongly disagree', 'disagree', 'neutral', 'agree', to 'strongly agree' were coded into numerical data of 1,2,3,4, and 5 respectively. For testing hypothesis A, answers were sorted into three groups based on reported a. age, excluding empty answers. Secondly, b. gender was grouped into a group of 'male', while combining 'female' and 'diverse'. Empty answers, or 'prefer not to say' were excluded. Combination of reported female and diverse gender in one group should not re-construct a binary conception of gender. Hypothesis B theorises gender as an aspect of vulnerability to data

breaches, assuming 'female' and 'diverse' participants to be more vulnerable to violations of online privacy (see Savoia *et al.*, 2021).

For analysis of hypothesis C, reported internet usage habits were grouped into three groups: 'every day', 'every day with paternal oversight', and 'rare use'. Participants reporting no use at all, or empty answers were omitted. Hypothesis D investigates whether preferences for ease of access online might account for a change in participants' general attitude towards online privacy. These attituded were measured in the 'nothing to hide' argument of question 19 or the 'data shared online is irrelevant' statement in question 16. A preference for ease of online access was determined by reporting 'yes' or 'no answer' to question 11 'I use a password manager on my mobile phone'; and question 12 'I enter different sites via my google or Facebook or similar account'. Only a definitive 'no' was grouped as clear preference of online privacy over ease of access.

Each hypothesis compared the average level of agreement towards a statement between two groups. Relevant group's means were then interpreted as general attitudes on the Likert scale. The sample maximum 'strongly agree' and sample minimum 'strongly disagree' was translated into numerical data of 5 and 1. The difference between the largest and smallest values is the range of the sample, which is (5 – 1) = 4 points of agreement. This means the interval for the Likert scale ratings is the range of 4 divided by total the total of the scale - hence, (4 / 5) = 0.80 data points. Applying this interval to the Likert scale allows for reverse translation of the numerical data into levels of agreement, laid down in following table (Table 3).

Mean Value	***Level of Agreement***
1.00- 1.80	**strongly disagree**
1.81-2.60	**disagree**
2.61-3.40	**neutral**
3.41-4.20	**agree**
4.21-5.00	**strongly agree**

Table 3. **Interpretation of Means of Grouped Answers.**

Given the small sample sizes of grouped answers, the change in mean levels of agreement between two groups was then tested for statistical significance with a two-tailed two-sample t-test. At $p > .05$ equal means of both populations need to be assumed ($\mu_1 = \mu_2$) and respective null hypotheses $H_{A\text{-}D0}$ need to be accepted. At $p \leq .05$ the null hypothesis is rejected and the alternative hypotheses $H_{A\text{-}D1}$ accepted. The means of two populations are not equal ($\mu_1 \neq \mu_2$) in this case, and a significant change in mean levels of agreement is to be observed between relevant groups.

V. Results

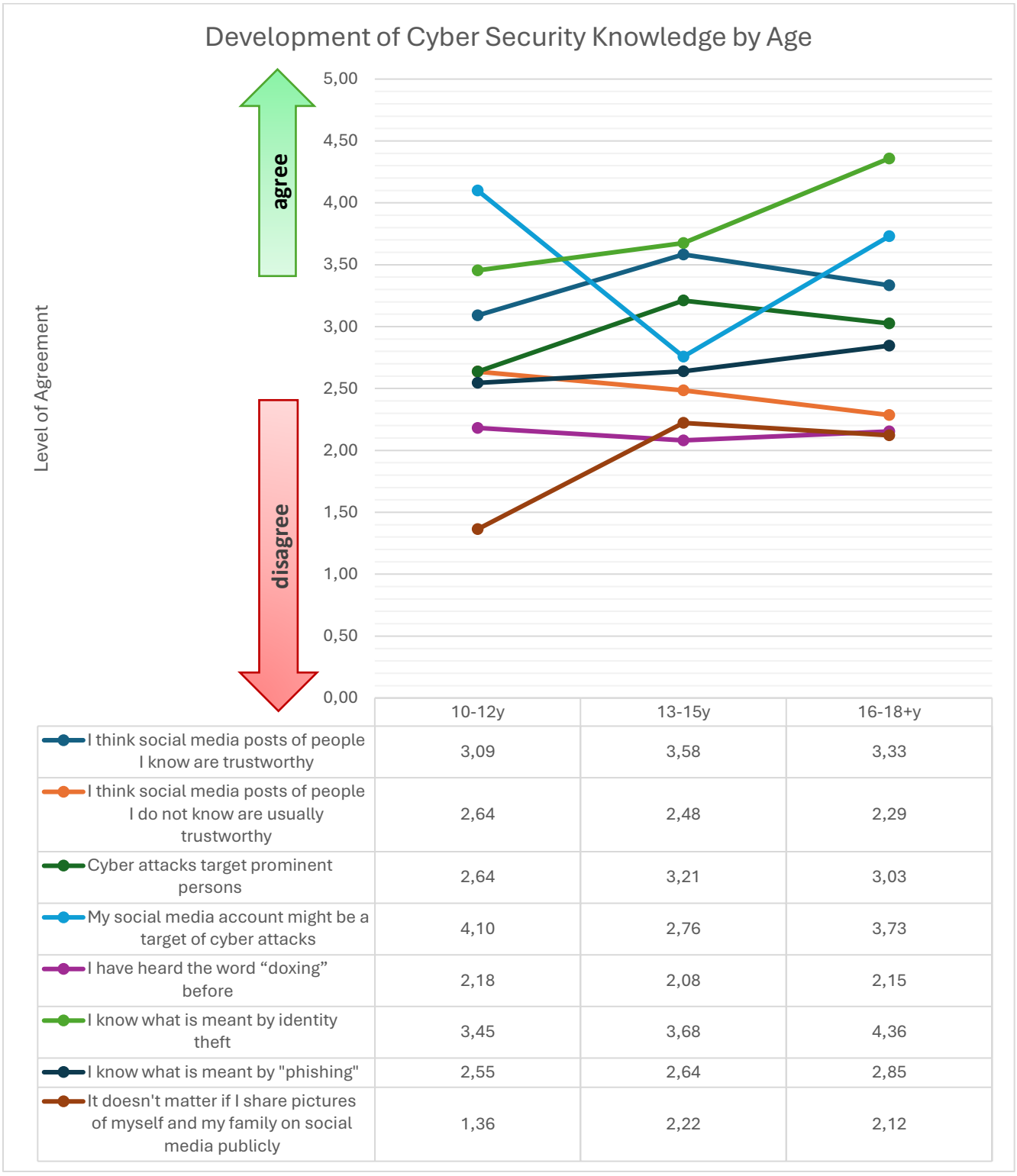
a. Hypothesis A. Differences in Data Protection Awareness based on Age

To estimate how age as a variable affects awareness of data protection issues, and their knowledge of cyber security threats, students were grouped in three age groups: the youngest group encompassing 10, 11, and 12 years, the middle group 13, 14, and 15 years, and the oldest group 16,17, 18 and 18+ years of age.

All questions required a level of agreement on a 5-point Likert scale towards statements like, for example 'I think social media posts of people I know are trustworthy', and 'I think social media posts of people I do not know are trustworthy' (see questions 6 A1 and A2 in Table 4 and 5). In this example, only a slight change in agreement to trust in known peoples' posts is visible between the youngest age group of 10–12-year-old students to 13-15 years old students. While earlier group is neutral towards the statement with M=3.09, latter rather agrees with M=3.58. Vice versa, distrust in strangers' posts seems to increase with age: while the youngest age is neutral towards trust in posts of unknown people with M= 2.64, the older age groups disagree with M= 2.48, and M= 2.29 (see Table 4). These changes are not statistically relevant.

Similar trends are to be observed regarding cybersecurity knowledge, although not all are statistically significant. The average agreement of knowing terms like 'identity theft', or 'phishing' progressively grows with age, as to be expected (Table 4). For example, the oldest age group strongly agree M=4.36 that they know what is mean by 'identity theft' in statement A2 of question 8, while their younger peers simply agree with M=3.67, and M=3.45. This difference is statistically significant with $t(85)=-2.99$, $p= .0004$. Another significant outlier with $t(79)=4.10$, $p= .0001$ is the 13–15-year-old participants' neutral stance M=2.75 towards statement A2 of question 7 'My social media account might be a target of cyber attacks', while their younger peers M=4.1, as well as their older peers M=3.73 agree.

Interesting outliers are *strong disagreement* with M=1.36 of the youngest age group to the statement A1 of question 16 'It doesn't matter if I share pictures of myself and my family on social media publicly', while their older peers *simply disagree* with M=2.22, and M=2.12. This is a statistically significant difference of $t(87)=-2.16$, $p = .03$. Both in cyber security and data protection awareness, H_{1A} may thus be supported: the mean level of awareness changes with the participant's age.



	10-12y	13-15y	16-18+y
I think social media posts of people I know are trustworthy	3,09	3,58	3,33
I think social media posts of people I do not know are usually trustworthy	2,64	2,48	2,29
Cyber attacks target prominent persons	2,64	3,21	3,03
My social media account might be a target of cyber attacks	4,10	2,76	3,73
I have heard the word "doxing" before	2,18	2,08	2,15
I know what is meant by identity theft	3,45	3,68	4,36
I know what is meant by "phishing"	2,55	2,64	2,85
It doesn't matter if I share pictures of myself and my family on social media publicly	1,36	2,22	2,12

Table 4. **Average Group Means for Hypothesis A.**

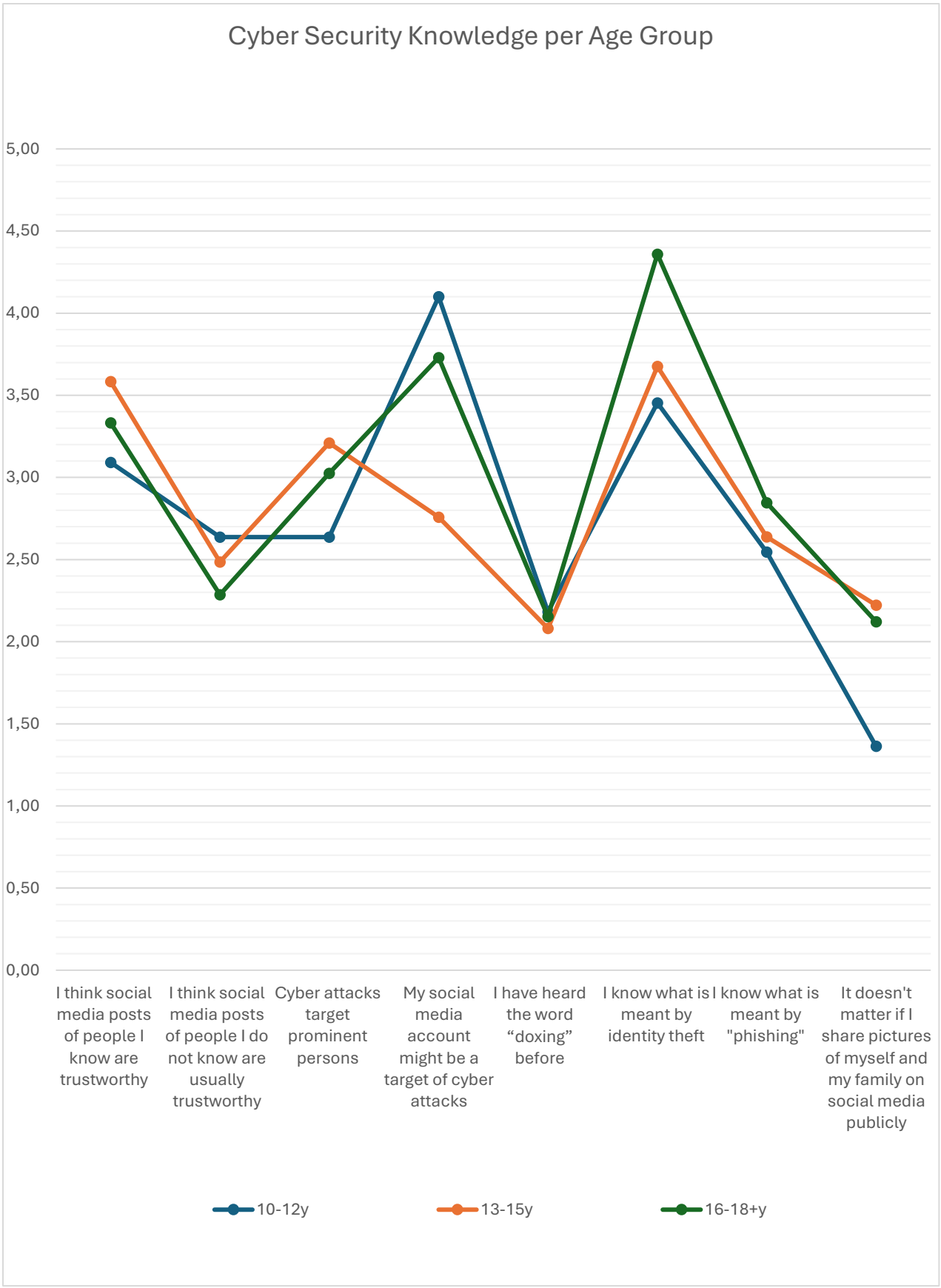


Table 5. **Change in Group Means per Statement for Hypothesis A.**

b. Hypothesis B. Gender differences in Online Trust, and Privacy Attitudes

Based on the research framework, the hypothesis of change in agreement levels to statements of privacy based on reported genders was tested. Question 19 operationalises data protection awareness, and attitudes towards online privacy in various statements the participants could agree or disagree to with a 5-point Likert scale. There are no significant differences in agreement levels towards the statements arguing it is too late to protect data online (A2), that formal requirements of data protection are impractical (A8), nor between the statement asking students whether they pro-actively implement data protection measures (A3). On average, students of any gender are neutral towards these statements. When it comes to distrust in companies to handle personal data responsibly (A4), female and diverse students are rather *neutral* towards the statement with on average M = 3.38 points of agreement, while male participants rather *agree* with M = 3.45 points of agreement. With the threshold value for agreement being at 3.41 average points of agreement, the difference is not statistically significant. Student of every gender agree that they are responsible for the protection of both their own and their families' and friends' personal data (A5).

There is a significant difference with $t(156)= 2.28$, $p= .024$ in agreement to the nothing to hide argument (A1). Female and diverse participants *disagree* (M=2.54, SD=1.18) to not being worried about their online data because they have 'nothing to hide' (A1), while male participants (M=2.98, SD=1.23) are neutral towards the statement. Significant difference with $t(145)= 2.06$, $p= .043$ is also to be observed in the participants' ignorance of the implications of online consent. Male participants (M= 2.97, SD=1.28) were *neutral* towards the statement 'Often I am just clicking around, not sure what I accept with the 'accept' button' (A7), female and diverse participants rather *agreed* (M=3.4, SD=1.18). Only for the 'nothing to hide' argument (question A1), and insecurity about the extent of notice and consent models online (question 19 A7) can therefore the alternative hypothesis H_{B1}: 'the mean level of attitudes towards online privacy changes with the participant's gender' be supported.

Other potential factors like cybersecurity knowledge, knowledge of privacy rights and trust in online environments were also tested. There was no

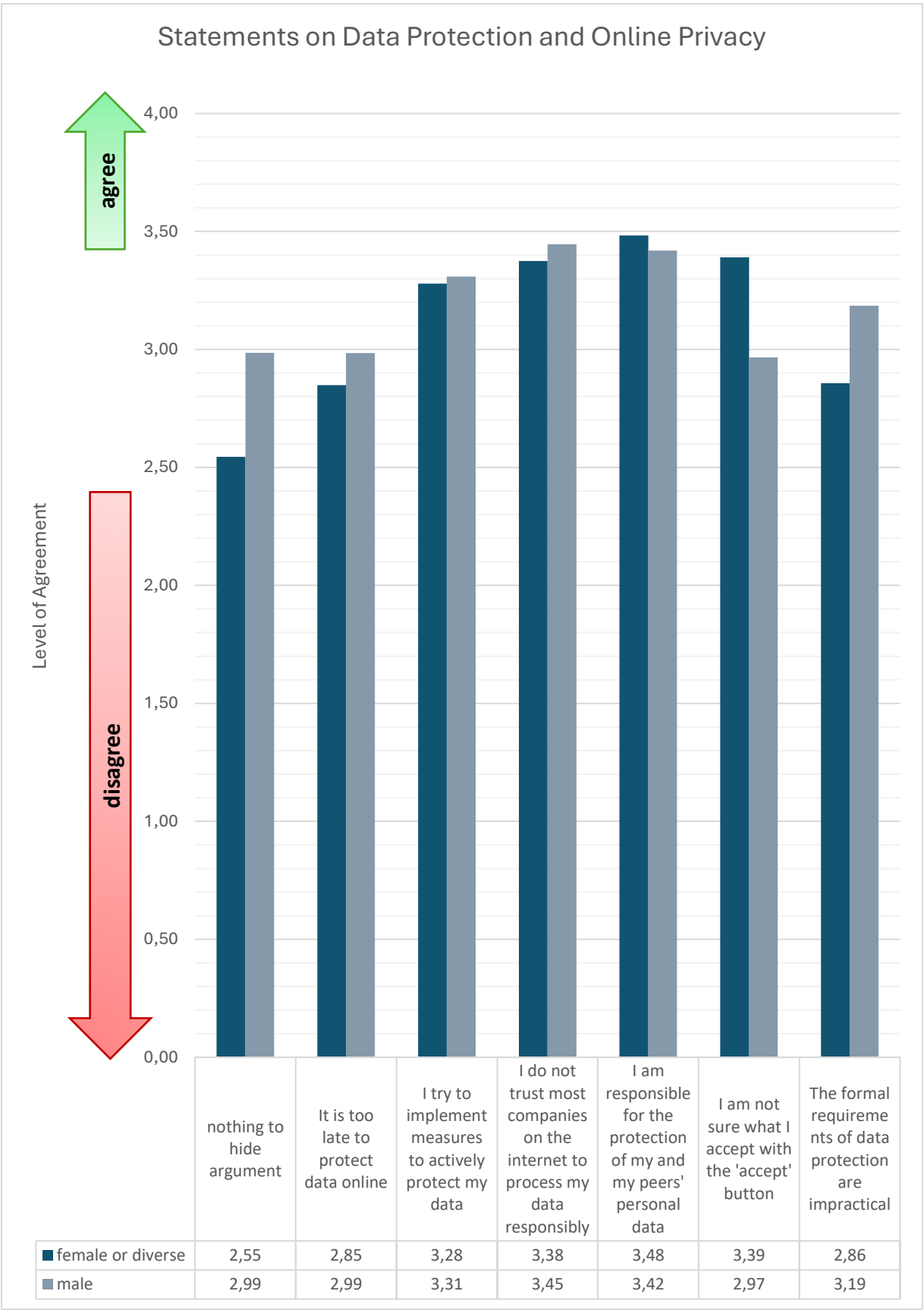


Table 6. **Change in Gendered Group Means per Statement for Hypothesis B.**

significant difference in agreement levels on these topics based on gender. In detail, there was no significant difference between genders in answers to whether posts of known, or unknown people on the internet are trustworthy (Question 6 A1-2). Male participants *agreed* M=3.66 like female and diverse participants M=3.41, that posts of known people are usually trustworthy (A1), Both groups *disagreed* to whether they trust posts of unknown persons (male groups M= 2.53; female or diverse group M=2.4).

Concerning the awareness of cybersecurity issues queried in question 7 A1 and A2, all genders were on average *neutral* about their, or famous persons' social media account being the target of a cyberattack. No difference in cyber security knowledge could be detected between genders, all *agreed* to know the terms doxing, identity theft, and phishing (see Question 8 A1-3).No significant difference in attitudes towards privacy as a right was found between the groups, all *agreed* that corporate forces have too much influence in Question 18 A1, while *agreeing strongly* with privacy as a right and responsibility of data protection (question 18 A 3 and A 4). All groups were *neutral* towards statements of governmental control (Question 18 A2). Thus, the null hypothesis H_{B0}: 'the mean level of trust in online environments does not change with the participant's gender' needs to be supported in these cases.

c. Hypothesis C. Differences in Data Protection Awareness based on Usage Habits

To estimate whether privacy and data protection attitudes relate to the participants' self-reported habits of online use and vice versa, the change of agreement levels towards Question 19 'Please rate the following statements in concern of data protection...' was mapped (Table 7). Participants reporting to use the internet 'Every day, as much as I want' were grouped into daily users (N_1=135 out of N=196), while their peers who use it every day, but under their parents' control, were analysed separately (N_2=37). Participants who reported very limited use, or no internet use at all were grouped together (N_3=17). The differing levels of agreement to statements provided in Q19 were analysed to test whether different usage habits account for change in level of awareness of data protection or privacy issues. For example, paternally or self-restricted use of the internet could point at sensibilisation for data protection, and for privacy issues.

The 'nothing to hide' argument in statement A1 revealed differences among the groups. While everyday users without restrictions were *neutral* towards this argument, users with paternal oversight and restrictive users rather *disagreed*. While the difference between the average agreement of the groups was not statistically significant, it could still indicate relations between restrictive use of the internet and the understanding of online privacy and the value of personal data.

Such insights correspond to the statistically significant difference with p = .0026 (t (120) =3.21, p<.05) in attitudes towards online data protection, being more positive for participants who have restricted internet usage. Restrictive users *disagreed* on an average of M=2.06 with statement A2 'We all leave so many traces on the internet, so it is too late to protect it' while non-restrictive users where rather neutral about the statement. The alternative hypothesis H_{C1}: 'mean levels of privacy attitudes change with online usage habits of the participant' needs to be accepted in the case of attitude towards data protection queried in statement A2.

However, when it comes to real-life application of data protection measures, there is no significant difference in behaviour. The statement A7 'Often I am just clicking around, not sure what I accept with the 'accept' button' was operationalised as the participants approach towards data protection. Disagreement could point at higher confidence in applying data protection measures, whereas agreement shows uncertainty regarding data protection. The average of every group's agreement indicates a *neutral* stance towards this statement, with no significant differences based on usage habits.

Similar observations can be made in participants' distrust in corporations: non-restrictive everyday users, everyday users with paternal oversight as well as restrictive users *agree* with statement A4 'I do not trust most companies on the internet to process my data responsibly', with no significant differences. It can therefore be argued that the participants are aware of the monetary value of their personal data, and internet usage habits do neither relate to trust nor distrust in online services. Similar applies to the awareness of responsibility for data protection and its formal requirements: these are *not* dependent on usage habits. No significant difference in agreement to statement A5 'I am responsible for the protection of my personal data, the data of my family members, friends and others', nor A8 'The formal requirements of data protection are impractical' were found. Overall, usage habits or paternal oversight seem to have only limited effects on the participant's awareness of (the value of) data protection, and the null hypothesis H_{C0}: 'mean levels of privacy attitudes do not change with online usage habits of the participant' is accepted.

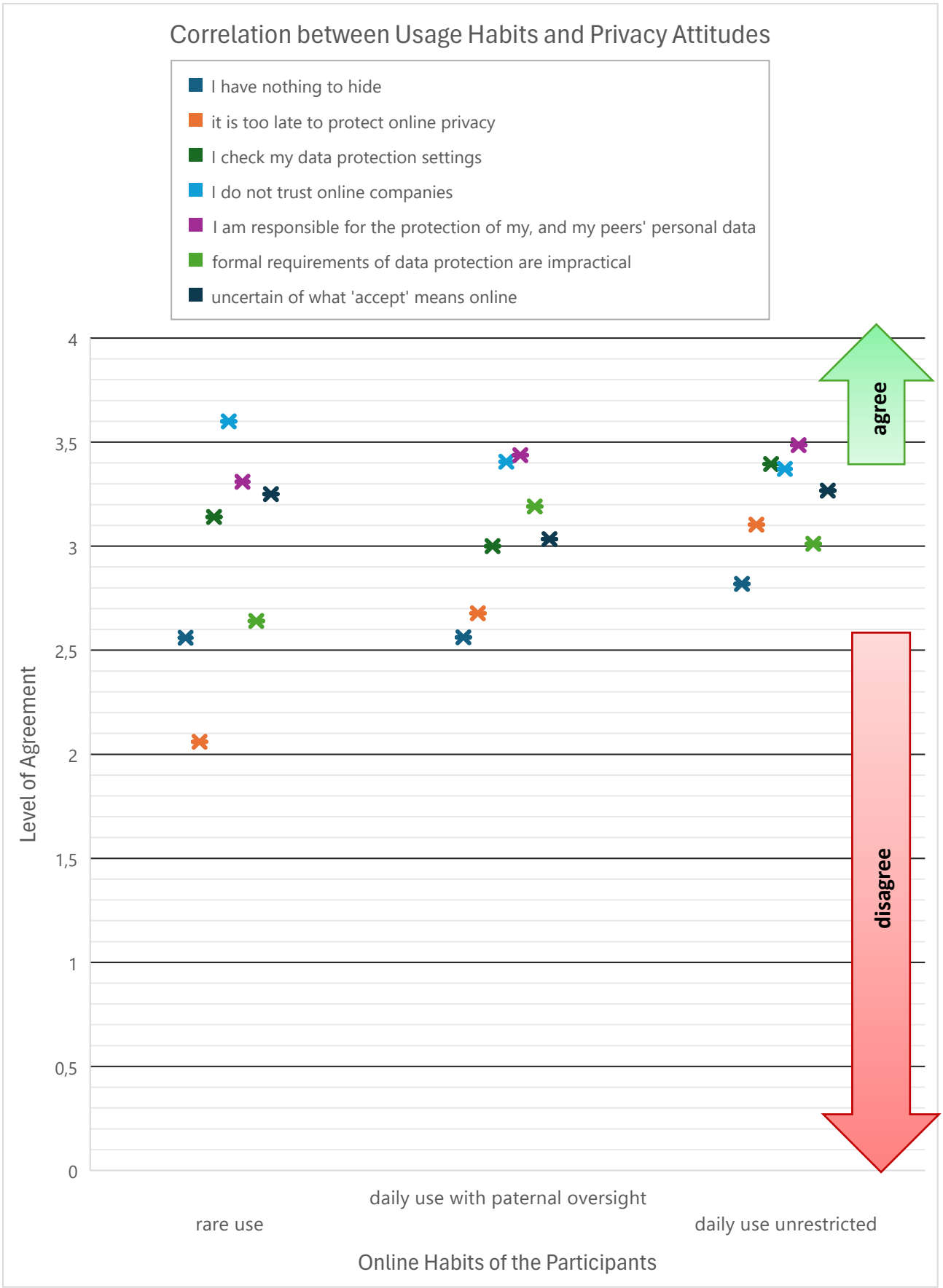


Table 7. **Change in Group Means per Statement for Hypothesis C.**

d. Hypothesis D. Differences in Privacy Attitudes based on Privacy Preferences

There was no significant change in agreement to statements on the value of personal data based on a loose attitude towards data protection (see Table 8). The survey participants' attitude was estimated by operationalizing their negative, positive, or neutral answer to Q11 'I use a password manager on my mobile phone'; and Q12 'I enter different sites via my google or Facebook or similar account.'

'Yes' and 'No answer' served as indicator for a rather loose attitude towards data protection, and favouring ease of access to sites over privacy. 'No' was operationalized as definitive preference of data protection over ease of access. Interestingly, students who use password managers on their mobile devices (see Q11) and have their google or Facebook account linked to several services (see Q12) do *not* necessarily agree to arguments like 'I have nothing to hide' (see question 19 A1), nor to sharing of personal data to be irrelevant (see question 16 A1).

A change in average levels of agreement from simple disagreement M=2,15 of the group which prefers ease of access, to *strong* disagreement M=1,78 of the group who prefers data protection is detectable, although not statistically significant with p .14 (Table 8). These findings lead to a rejection of the alternative hypothesis H_{D1}, and acceptance of H_{D0}: 'mean levels of privacy attitudes do not change with the participant's online preferences.' From the sample analysis, it cannot be inferred whether participants who prefer ease of access with password managers and linked accounts are less aware of privacy issues.

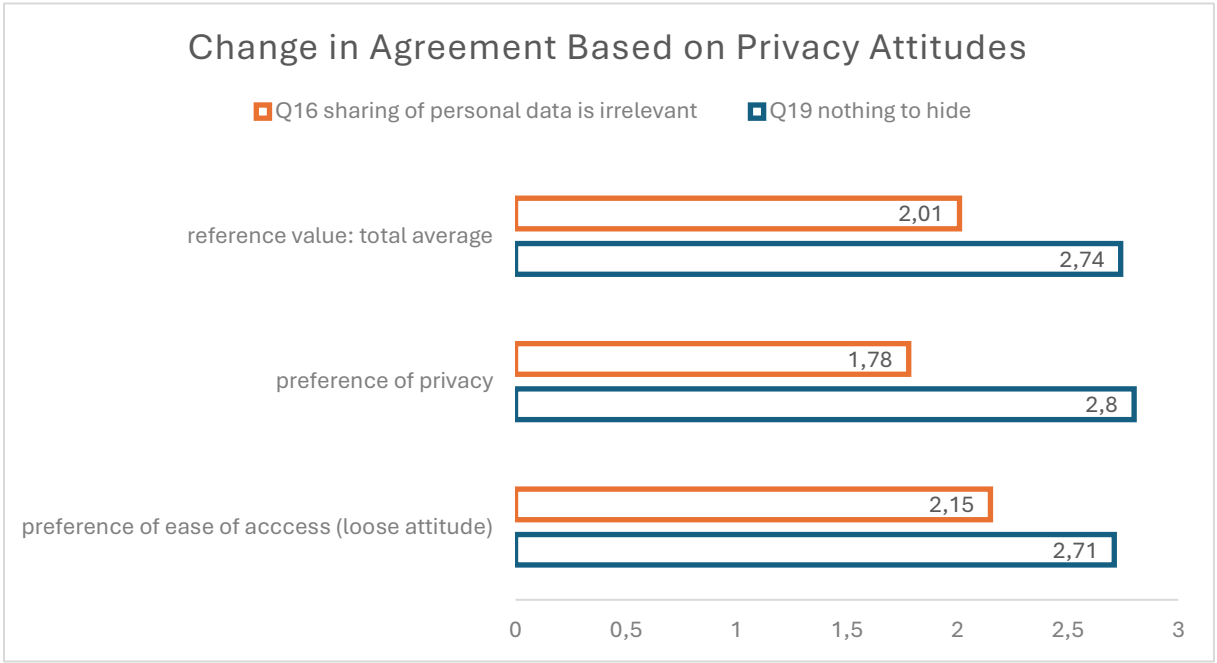


Table 8. **Change in Group Means per Statement for Hypothesis D.**

VI. Discussion and Conclusion

This study investigated factors influencing students' awareness and attitudes towards data protection and online privacy, Analysis focused on age, gender, internet usage habits, and online privacy preferences as potential factors. The findings illuminate how the students' understanding of data protection may change according to demographic variables or online habits and preferences. A concise summary of significant changes is given in following chapters, interpreting the results and connecting them to former research.

a. Age is a Factor in Students' Level of Data Protection Awareness

Hypothesis A explored the relationship between age and cybersecurity awareness. While there is no significant change in agreement towards questions of trust in the online environment, significant change between groups was detected in their cyber security knowledge. The older participants significantly agreed more to know terms of cyber security like 'phishing' (see Table 4).

Interestingly, the group pf 13-15 years old participants were significantly more indifferent to whether their own social media account might be a target of cyber attacks'. Both the younger and older age group agreed to this statement. Outlier like these in the data shall not be overinterpreted due to the small sample size. These significant differences may still be starting point for further discussion, for example on the effect of the Covid19-pandemic on early cybersecurity education. Respective group might have been in the critical age of 9 to 11 years when the pandemic started, meaning they may have experienced transition from elementary to secondary school from home.

Another significant outlier in the sample analysis was the youngest group's *strong disagreement* to: 'It doesn't matter if I share pictures of myself and my family on social media publicly'. Their older peers *simply disagreed*. Although cyber security and data protection knowledge seems to rise with age, the attitude towards online privacy seems to decline. Being allowed to open accounts on most social media services at the age of 13 may account for such downward trend. Another explanation might be more exposure to online content with higher age groups, leading to indifference.

b. Minor Impact of Gender on Students' Attitude towards Online Privacy

The study examined gender differences in online trust and privacy attitudes using self-reported agreement levels on various statements. Most aspects of trust and awareness showed no significant gender-based differences, which corresponds to the theory that the more independent use of the online technology happens on a day-to-day basis, the less it is used to perform a certain gender identity. Rather than casual occurrence, the use of social media, and navigation of the online realm as

a routine makes it ‘inapt as a resource for gender differentiation’ (Remmele and Holthaus, 2013, pp. 28, 39).

At the same time, some notable exceptions emerged. Female and diverse participants disagreed more than males with the ‘I have nothing to hide’ – statement. The ‘nothing to hide’ argument has been criticized for assuming the inherent value of privacy is bound to lawfulness of behaviour and there would be 'no threat to privacy unless the government uncovers unlawful activity, in which case the person has no legitimate justification to claim it’ (Solove, 2007, p. 746). The conflation of individual privacy with threats to public security becomes necessary for this assumption to work. Such de-valuation of individual privacy enforces power structures between people and institutions and create a sense of helplessness in the individual (2007, p. 757). Hence, the female and diverse group’s stronger stance towards the ‘nothing to hide’ argument might be interpreted as sense of higher vulnerability in such power structures. Former research pointed out that young female internet users would be more likely to ‘encounter specific risky situations in the online space (…) independent of other online behaviors such as the amount of time spent online or type of social media platform being used’ (Savoia *et al.*, 2021, p. 12).

c. Data Protection Awareness does not change with Internet Usage Habits

Scarce internet usage revealed significant change in agreement to statements that prioritize data protection. Restrictive users disagreed to question 19 A2 ‘We all leave so many traces on the internet, so it is too late to protect it’ and thereby showed an assertive stance on data protection. Similar was to be observed for the ‘nothing to hide’ argument. Nevertheless, no significant differences were observed in practical measures of data protection, such as responses to online consent or distrust in corporations. Restrictive use or paternal supervision might therefore not affect online behaviour, while it does seem that less or supervised internet use fosters greater conceptual awareness of data protection.

Tool development should therefore consider that frequent users need to be especially sensitised for the value of their privacy beyond having ‘nothing to hide’. Resignation towards data protection as visible in the statement ‘it is too late to protect’ data online is another relevant difference between frequent and rare users. If more frequent use corresponds to a subjective sense of uselessness in data protection measures, education should not only include the ‘how’ of protecting data, but also the ‘why’.

d. Preference of Ease of Access does not equal a Loose Attitude towards Privacy

Online preferences were analysed with reported behaviours such as using password managers or linking social media accounts to other services. Participants not exhibiting such behaviour strongly disagreed with statements supporting a loose attitude toward data sharing, such as “I’m not worried about my data because I have nothing to hide.” However, so did their peers who preferred ease of access by for example, using password managers. Slight differences in disagreement to the indifference of sharing data online were not statistically significant, leading to acceptance of the null hypothesis (D0)

Online preferences were analysed with reported behaviours such as using password managers or linking social media accounts to other services. Participants not exhibiting such behaviour strongly disagreed with statements supporting a loose attitude toward data sharing, such as “I’m not worried about my data because I have nothing to hide.” However, so did their peers who preferred ease of access by for example, using password managers.

Slight differences in disagreement to the indifference of sharing data online were found among the two group: definite preference for online privacy corresponded with strong disagreement, while preference for ease of access with simple disagreement. This difference was not statistically significant, leading to acceptance of the null hypothesis (D0) (compare to Table 8).

Therefore, it can be argued that students who share their data more frequently with online services to ease access and navigation might not necessarily be indifferent about their online privacy. Rather, online behaviour that compromises online privacy seems independent from the actual awareness of the value of privacy and data protection.

e. Identified Learning Hurdles

In conclusion, for development of user-based learning tools, *learning hurdles* may be inferred from the tested hypotheses:

A. Young age might influence less awareness of cyber security threats, while higher age groups might get de-sensitized for data protection issues due to higher exposure. This hurdle may be tackled with tools sensitive to the user's needs in specific age groups.

B. Gender does not represent a learning hurdle for data protection per se but might be a factor for higher vulnerability to privacy breaches.

C. Everyday use of the internet is associated with less awareness of the value of online privacy. However, usage habits do not affect knowledge on concrete data protection measures. The 'why' of data protection would need to be stressed to address this hurdle.

D. Paradoxically, users who share data more frequently and voluntarily do not value online privacy less. Rather, ease of navigation and access is an important factor for the online experience of students. A learning hurdle is for example, the habitual linking of services, which may be overcome with offering students outlook on privacy-friendly alternatives.

VII. References

a. Figures

b. Tables

c. Bibliography

Nagel, J. (2024) 'LimeSurvey - Free Online Survey Tool.' Hamburg: LimeSurvey GmbH. Survey Services & Consulting. Available at: https://www.limesurvey.org/ (Accessed: 01 December 2024).

Remmele, B. and Holthaus, M. (2013) 'De-gendering in the use of e-learning', *The International Review of Research in Open and Distributed Learning*, 14(3), pp. 27–42. Available at: https://doi.org/10.19173/irrodl.v14i3.1299.

Remmele, B. and Valentic, Z. (2024) 'Curriculum on Data Protection'. DataPro Project.

Savoia, E. *et al.* (2021) 'Adolescents' Exposure to Online Risks: Gender Disparities and Vulnerabilities Related to Online Behaviors', *International Journal of Environmental Research and Public Health*, 18(11), p. 5786. Available at: https://doi.org/10.3390/ijerph18115786.

Solove, D.J. (2007) '"I've Got Nothing to Hide" and Other Misunderstandings of Privacy'. Rochester, NY: Social Science Research Network. Available at: https://papers.ssrn.com/abstract=998565 (Accessed: 23 December 2024).

Valentic, Z. and Strauß, H. (2024) 'First Results of DataPro Survey. Data Protection between Individual Agency, Surveillance and Making Money.' DataPro Project.

VIII. Appendix

A. LimeSurvey Detailed List

Respective Variable	Survey Question	Question Code
Demographic Data	How old are you) (multiple choice)	Q20
	Are you (female/male/diverse/prefer not to say)	Q21
	Are you from… (Multiple choice)	Q22
Internet Usage Habits	How many social-media accounts do you have? (multiple choice)	Q02
	How often do you use the Internet? (every day/ very day, but my parent control it/ only a few times a week/ I do not use it at all)	Q03
	If you have social media... (I have a public profile/ I have a private profile)	Q04
	Which social media platform do you use most? (multiple choice)	Q05
Data Protection Awareness		
Data Protection Skills	How does a safe password look like? (Multiple choice, right/wrong/no knowledge)	Q10.SQ1-SQ5
	I use a password manager on my mobile phone. (yes/no/ do not know)	Q11
	I enter different sites via my google or facebook or similar account. (yes/no/ do not know)	Q12

	I check and customize the data protection settings on my mobile device. (yes/no/ do not know)	Q13
Cyber Security Knowledge	I can detect a phishing attempt by … (Multiple choice, right/wrong)	Q09.SQ1-SQ5
	Cyber attacks target prominent persons. (Likert scale rating)	Q07A1
	My social media account might be a target of cyber attacks. (Likert scale rating)	Q07A2
	I have heard the word “doxing” before. (Likert scale rating)	Q08.SQ1
	I know what is meant by identity theft. (Likert scale rating)	Q08.SQ2
	I know what is meant by "phishing". (Likert scale rating)	Q08SQ3
Privacy Rights Awareness		
Knowledge on Legal Rights	I have the legal right (multiple choice) - to obtain information about stored data - to correct information about others	Q15 SQ1 SQ2

	- correction of the data stored about me	SQ3
	- deletion of data stored about me	SQ4
	- that my data is stored until I delete it in 30 years	SQ5
	- restriction of data processing, if not required	SQ6
	- skip	SQ7
Trust in Online Environments	I think social media posts of people I know are trustworthy.	Q6.A1
	I think social media posts of people I do not know are trustworthy.	Q6.A2
Perception of Public-Private Dichotomies	Can you give some examples of personal data? (open-text)	Q14
Awareness of Value of Online Privacy	Personal data should be protected, because (multiple choice)	Q18
	- otherwise big companies could influence our behaviour	A1
	- otherwise the state could control us	A2
	- privacy is a basis for civil rights	A3
	- data protection is part of responsible behaviour	A4

	Please rate the following statements in concern of data protection... (Likert scale rating)	Q19
	- I'm not worried about my online data; I have nothing to hide.	A1
	- We all leave so many traces in the internet, so it is too late to protect it.	A2
	- I try to implement measures to actively protect my data., e.g. requesting and checking data information about me.	A3
	- I do not trust most companies on the internet to process my data responsibly.	A4
	- I am responsible for the protection of my personal data, the data of my family members, friends and others.	A5
	- Often, I am just clicking around, 0t sure what I accept with the 'accept' button.	A7
	- The formal requirements of data protection are impractical.	A8
	It doesn't matter if I share pictures of myself and my family on social media publicly. (Likert scale rating)	Q16.A1
Data Value Awareness		

Knowledge on Data Processing	Browser cookies ("pop-ups") track/save/ delete ... (Multiple choice, right/wrong/skip)	Q17
	I can detect so called 'dark patterns' which are used to induce unintended behaviour (Likert Scale rating)	Q16.A2