Development of a German Instrument for Self-Perceived Data Literacy

An Algorithm-based Approach to Scale Development

Leonie Hagitte

2024-05-26

Contents

Abstract

The increasing relevance of competent and critical handling of data in society not only makes it possible to record this competence, but also makes self-perception with regard to this competence increasingly clear. Previous approaches consider this competence primarily against the specific background of individual target groups, jobs or roles (Cui et al., 2023). In addition, only a few explicitly refer to the general population (Carmi et al., 2020; Cui et al., 2023). In view of the various theoretical approaches, there is a need for a uniform definition of data literacy in order to create comparability. Our aim is therefore to derive a holistic definition based on these approaches and to develop a questionnaire for self-perception of one's own data literacy. To this end, the decisive factors for the construct from previous definitions and operationalizations in various disciplines are brought together. Cognitive interviews are conducted iteratively to create and refine the items. The items are then selected using algorithm-based item selection. The facets of data literacy are comprehensively tested for factorial, discriminant, convergent and congruent incremental validity in order to promote a differentiated understanding of the construct. Construct and criterion validity are tested using correlations and hierarchical regression analyses, while cross-validation checks the robustness of the instrument. Based on a cross-sectional online questionnaire study, we first examine a representative sample of people from the general population. Limitations arise from the cross-sectional design and the heuristic item reduction, which limit predictions of predictive validity. The heterogeneous nature of the construct makes global instrument development and understanding of all participants difficult. The self-assessment questionnaire promotes a holistic assessment of competence and its perception for further research, for example by comparing self-assessment and actual performance.

Acknowledgements

1 Introduction

The relevance of data literacy in today's society becomes evident as it serves as a potent tool in navigating the complex data-driven environment. In a world characterized by information overload and rapid technological advancements, individuals equipped with strong data literacy skills can discern patterns, critically evaluate information, and make informed decisions

The exploration of citizens' interaction with media and the cultivation of their agency has traditionally centered around concepts such as written literacy, media literacy, information literacy, and digital literacy. In more recent discussions, Data Literacy has been approaching relevance among discussed competencies regarding what is necessary for agency in the current society (Carmi et al., 2020). Deficiency in data literacy not only exposes individuals to various risks and harms on personal, social, physical, and financial levels but also constrains their capacity to actively engage as informed citizens within an evolving, data-driven society (Carmi et al., 2020). Thus, Data Literacy is a competency that is becoming increasingly important to everyone. And research has acknowledged this in recent years, as more and more research is being done in that direction (Cui et al., 2023).

2 Background

Data Literacy involves the ability to effectively collect, manage, evaluate, and apply data in a critical manner. According to Wolff et al. (2016), it means being able to ask and answer everyday questions using both small and large datasets while considering ethical aspects. This includes skills such as selecting, cleaning, analyzing, visualizing, criticizing, and interpreting data, as well as communicating insights from data and using data for various purposes. (Frank, 2016) distinguish between cognitive skills, like data collection and analysis, and social skills, which involve trusting data while maintaining skepticism. (Calzada Prado & Marzal, 2013) outline five dimensions of Data Literacy: understanding data, acquiring data, interpreting and evaluating data, managing data, and using data. Understanding data includes knowing their types, roles, and significance, while acquiring data involves evaluating and selecting sources. Interpreting and evaluating data encompass understanding different presentation methods and data interpretation. Using data involves preparation, analysis, communication, and ethical considerations. Managing data includes storage, management, and reuse. This highlights one prominent feature of Data Literacy - It is a heterogeneous con-

cept. Every subject or profession seems to hold their own definition or framework of Data Literacy (Cui et al., 2023). While that most certainly is good for assessing specific skills (e.g. in an Recruitment test), it limits the generalisability and comparability of Data Literacy across individuals with different background. It furthermore limits the accuracy of communication about the topic as two people with different background might hold different definitions on Data Literacy. In the study from (Cui et al., 2023) it also becomes apparent, that one group seems to be underrepresented in the research on Data Literacy: citizens or the general public. While there seem to be several definitions of Data literacy for citizens (Wolff et al., 2016), most studies focus on other groups of people. This is the case although, the general public is forming the largest group by far, compared to professions like researchers, librarians, students or education provider (Cui et al., 2023). Thus, one might arrive at the question, why that ist? When taking a closer look into the concept and its comprising factors,it becomes clear that many of those factors tend to speak to professionals rather than laypeople. In their Framework (schüller, 2020) highlight the different roles in Data Literacy: Some of the facets or skills regard "data-consumption", whereas the most are skills "data-producers" would have. This is also reflected when taking a look into related concepts. The definition proposed by Wolff et al. (2016) suggests that data literacy shares some common competencies with statistical and information literacies.

Information literacy, often studied in library sciences, overlaps with data literacy in terms of accessing, critically evaluating, and using data sources (Calzada Prado & Marzal, 2013; Shields, 2005), Wolff et al. (2016) also emphasize the importance of the data inquiry process, starting from identifying problems, designing studies, acquiring data, conducting analysis, to drawing data-based conclusions. In comparison, Gould (2017) argued that data literacy is essentially the same as statistical literacy but with additional competencies needed due to the increasing importance of data. These added competencies include understanding who collects the data, how and why data is collected, and understanding data privacy and ownership (Gould, 2017). However, while statistical literacy focuses on quantitative data and basic statistics, data literacy extends to the ability to understand, access, evaluate, and use arguments and decision-making based on both quantitative and qualitative data (Cui et al., 2023).

2.1 Data and Information

With the closeness of those topics, it is worth while to take a look into the terminology, before delving deeper into the topic. When speaking of Data Literacy, one naturally has to think about what data is. What is information, how are they

different to each other and what extents do they share? And what role does the process of interpretation play? Does one interpret data to make sense of it? And if so, does that mean that data is entropy while information stands in opposition to it? The concepts of data and information are foundational in various fields, yet their precise definitions and relationships are often subject to interpretation. According to Shannon's seminal work on information theory (Shannon, 1948), data can be understood as raw, unprocessed symbols or observations, devoid of inherent meaning. It is through a process of interpretation and organization that data transforms into information, as elucidated by Bates (2005). Bates emphasizes that information emerges when data is structured and presented in a way that is comprehensible and relevant to a particular context or purpose.

Entropy, a concept borrowed from thermodynamics and applied in information theory, plays a crucial role in understanding the relationship between data and information. In his landmark paper, Shannon (1948) defines entropy as a measure of uncertainty or disorder in a system. In the realm of information theory, entropy is often associated with the amount of unpredictability or randomness in a set of data. However, it's essential to note that entropy can also be viewed as a measure of information content within a system. This perspective is articulated by Brillouin (1953), who suggests that low entropy corresponds to a high concentration of meaningful information. Similarly, the work of Jaynes (1957) highlights the connection between entropy and information, proposing that information can be quantified in terms of the reduction of uncertainty or entropy in a system. Thus, we can refine our understanding of the relationship between data, entropy, and information. While data serves as the raw material from which information is derived, it's the reduction of entropy through organization and interpretation that gives rise to meaningful information. Thus, rather than viewing data as synonymous with entropy or information, it's more accurate to consider information as emerging from the structured representation of data, leading to a deeper understanding of the underlying phenomena.

The framework of (schüller, 2020) shows that citizens are mainly covering roles where they consume data/ data products or informations. Thus, they are most likely to find tasks or items regarding the producing facets like providing or exploiting data more difficult to answer. This would systematically impair the fairness of tests and questionnaires regarding Data Literacy. Kubinger & Proyer (2005) define fairness as the condition where test measurements or values do not discriminate against specific groups of people who are relevant to the test. In the case of Data Literacy it is to be expected, that the item-difficulties in the producing factors would not be evenly distributed. As you cant just leave any factors out of the construct, in order to create a concept that is fair, it would be neccessary

to find a way to have the same factors in the concept with the same distribution of difficulty for all groups of participants.

2.2 Conceptual Integration

Thus, via conceptual integration we tried to find common grounds of the existing theories and studies. By synthesizing heterogeneous definitions and perspectives (as e.g. highlighted in the literature review from Cui et al. (2023)) into cohesive, unified representations and selectively incorporating relevant features while establishing cross-disciplinary connections. With questions in mind like "What factors and facets do most definitions share?", "What are specifics to certain definitions, addressed to special professions?", we arrived at the following structure of the concept:

Data literacy is the ability to collect, manage, evaluate, and apply data effectively. It involves asking and answering real-world questions from datasets while considering ethical use. Core skills include selecting, cleaning, analyzing, visualizing, presenting, critiquing, and interpreting data, information and their sources. We arrived at five core facets that are also very prominent in the most definitions in the literature (Cui et al., 2023). We further divided them into "consumer" facets (Comprehension, Evaluation & Integration), that are relevant for nearly every person in society, from citizens up. And "producer" facets (Communication & Statistics), that are mainly relevant for people, actively working with data.

2.2.1 Comprehension

This factor encompasses skills related to understanding and critically evaluating data and information. It involves the ability to comprehend various forms of data presentation, detect inconsistencies, interpret data comprehensively, and identify logical fallacies. Individuals with high scores on this factor demonstrate a strong aptitude for processing and making sense of complex information across different formats, enabling them to draw accurate conclusions and insights.

2.2.2 Evaluation

This factor involves skills related to critically evaluating information sources and discerning between facts and opinions. It encompasses the ability to assess the credibility and reliability of information, considering factors such as the reputation of the source and the context in which the information was presented. In-

dividuals scoring high on this factor demonstrate a keen awareness of potential biases or vested interests in information sources.

2.2.3   Integration

This factor relates to the ability to integrate data-driven insights into one's worldview and values. It involves actively seeking comprehensive understanding of various topics, engaging with diverse perspectives, and consciously incorporating data-driven insights. Individuals high in this factor adapt their opinions based on new data, prefer evidence-based information, and ensure their values align with reliable data. They engage with information and perspectives that challenge their existing views, showing a willingness to reassess their opinions and positions based on new data.

2.2.4   Communication

This factor revolves around the skill to effectively communicate and present data through various means, including visual formats, verbal explanations, and written descriptions. It requires translating complex data into clear and impactful formats, ensuring comprehension by varied audiences. Proficiency in data communication is essential for facilitating understanding, aiding informed decision-making, and prompting action based on data-driven insights. This proficiency includes the ability to translate data into simple visualizations, present findings confidently, and articulate complex information effectively in written and visual as well as verbal formats. It involves adeptly summarizing extensive datasets, engaging in professional discussions, and using advanced visual elements to convey specialized results to target audiences.

2.2.5   Statistics

This factor covers skills related to managing and analyzing data effectively. It involves proficiency in organizing and analyzing data using software tools, conducting statistical analyses, and understanding research methodologies. It includes skills such as organizing and managing data using software tools, conducting interviews or surveys for data collection, performing basic statistical analysis and recognizing trends in graphical representations. Individuals scoring high on this factor exhibit competence in statistical methods, enabling them to effectively analyze data and interpret findings.

2.3   Delineation from other concepts - The nomological net

In line with the framework of schüller (2020) the factors Data Comprehension and Interpretation, Contextualize and Evaluate Data and Data-Informed Worldview are representing the application areas for "consumers". So they are relevant for all people in scociety, whereas the other factors, those speaking to people who are considered "producers", are only common for smaller groups of people or certain professions.
This structure encompasses certain characteristics and behaviors from critical thinking, media competency, technology competency, statistical Literacy as well as from information literacy.

2.3.1   Statistical Literacy

Statistical literacy, as defined by Gal (2002), encompasses five knowledge dimensions: literacy skills, statistical knowledge, mathematical knowledge, contextual knowledge, and critical reflection skills, along with two dispositional dimensions: beliefs and attitudes, and a critical mindset. It involves understanding written, spoken, or graphical information, knowing why data are needed, grasping basic statistical concepts, understanding probability, and drawing statistical inferences. It also includes basic mathematical operations, interpreting information within a context, and critically questioning the validity of information, particularly in media. Statistical literacy requires a critical attitude and the perception of oneself as competent in statistical reasoning.

2.3.2   Information Literacy

Information literacy is the ability to recognize when information is needed and to effectively locate, evaluate, and use that information (American Library Association (2000). Information Literacy Competency Standards for Higher Education. http://www.ala.org/acrl/standards/informationliteracycompetency). It empowers individuals to determine the extent of information needed, access it efficiently, critically evaluate information and their sources, incorporate information into their knowledge base, and use it effectively for specific purposes. Additionally, information literacy involves understanding the legal, economic, and social aspects of information use and ensuring ethical and legal access and use of information. Furthermore, Information literacy is not static; it involves a commitment to lifelong learning and adaptation to evolving information technologies and practices. This definition already highlights the closeness of Information literacy to Data Literacy. Another possible way to describe it is suggested by Johnson and

Webber (2003): "Information literacy is the adoption of appropriate information behaviour to obtain,through whatever channel or medium, information well fitted to information needs,together with critical awareness of the importance of wise and ethical use of information in society." (Johnston & Webber, 2003, p.336)

2.3.3 Critical Thinking

The conceptualization of Critical Thinking (CrT) has evolved along three main branches: philosophical, psychological, and educational (Rear, 2019). In the philosophical view, which centers on the mental process of thought, a critical thinker is someone adept at logically evaluating and questioning both the assumptions of others and their own. On the psychological front, which delves into the processes driving action, a critical thinker possesses a combination of skills enabling them to assess a situation and determine the most appropriate course of action. The educational approach aligns more closely with the psychological perspective, relying on frameworks and learning activities tailored to enhance students' CrT skills and subsequently assess their proficiency in these skills (Payan Carreira et al., 2022).

2.3.4 Information and Communication Technology Competency

Information and Communication Technology Competency includes general and domain-specific ICT competencies, including communication, processing and storing, content generation, safe application, and problem-solving skills.

Our definition incorporates statistical literacy by emphasizing data interpretation, analysis, and understanding different types of data representations, such as graphs and tables (Data Handling and Analysis). It includes elements of information literacy by focusing on evaluating the credibility of data sources, considering factors like reputation and biases, which are similar to assessing the quality of information sources (Contextualize and Evaluate Data). Both statistical and information literacy involve using data and information to make informed decisions. Data literacy emphasizes integrating data-driven insights into ones opinions and values, that then later on influence decision-making processes, aligning with the goals of statistical and information literacy. The factor "Data Comprehension and Interpretation" encompasses behaviors and skills associated with critical thinking, such as the capacity to identify weaknesses in one's reasoning or to actively shape discourse and, consequently, the public dissemination of information. The factor "Contextualize and Evaluate Data", as well as the "Statistics" factor, encompass behaviors and skills related to technology competency. This includes

abilities such as navigating and critically evaluating online sources and platforms, using information and communication technology, as well as utilizing statistical software, among others. The "consumer" factors are also closely related to media literacy, which focuses on skills related to critically analyzing and interpreting various media formats, including text, images, and videos, for understanding and engaging with media content.

In contrast, statistical literacy often focuses more narrowly on statistical concepts and methods, such as probability, sampling, and hypothesis testing. This definition encompasses a broader range of skills beyond statistical concepts, such as data visualization, software usage, and understanding data collection methods. While statistical literacy often involves analyzing existing data, this definition also encompasses skills related to producing and manipulating data. This includes tasks such as creating visualizations and using software tools. This aspect goes beyond traditional statistical literacy and aligns more closely with data literacy. While information literacy involves assessing the quality of information sources, this definition places a particular emphasis on assessing data quality, considering factors like sample size, biases, and data context. This aspect extends beyond traditional information literacy and is more specific to data literacy. To contrast data literacy from technology competency, it is to be highlighted, that data literacy involves proficiency in using technology, but thereby focuses specifically on understanding and working with data. Technology competency encompasses a broader set of digital skills that extend beyond those relevant for data literacy.
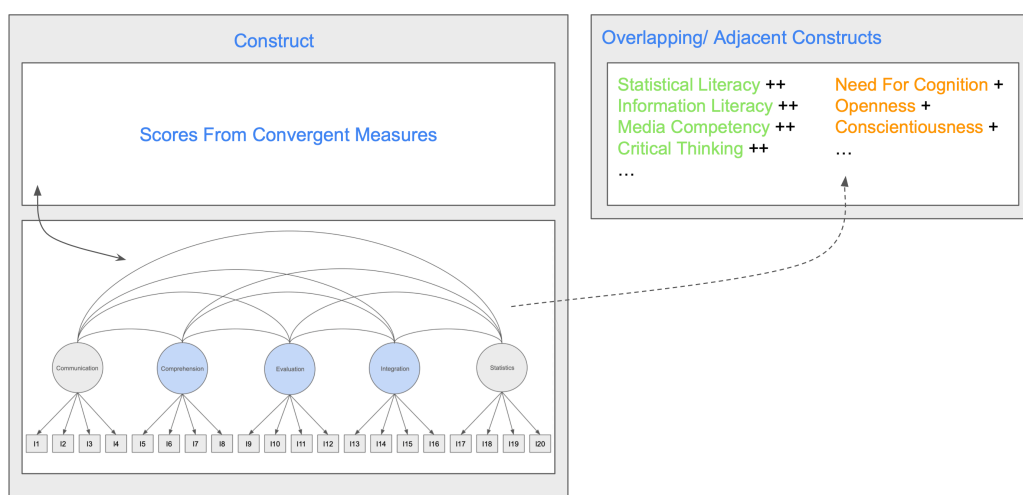

Figure 1: Illustration of the nomological net of data literacy.

2.4 Discrimminant Constructs

2.4.1 Need for Cognition

The personality trait known as Need for Cognition (NFC) originated in social psychology during the 1940s and 1950s, the concept of NFC, representing an inclination for joyful thinking, is evident in the works of Maslow (1943), Murphy (1947), Asch (1952), and Sarnoff & Katz (1954). The conceptualization of NFC underwent refinement in the mid-1950s through experimental investigations by Cohen and colleagues (Cohen et al., 1955). They defined NFC as "a need to structure relevant situations in meaningful, integrated ways. It is a need to understand and make reasonable the experiential world" (Cohen et al., 1955, p. 291). The concept captures individual variations in the engagement and enjoyment of thinking tasks (Bless et al., 1994).

2.4.2 Opennes to new experiences (B5)

2.5 Aim of the Study

My aim was to derive a comprehensive definition of data literacy based on existing approaches and to develop a questionnaire for self-perceived data literacy, measuring the three core factors of the construct. While drafts for two additional factors are included as preliminary assessments for future studies, they are not the focus of this study.

3 Methods

3.1 2. Stage Two: Item Selection and Construct Validity

- Conduct a quantitative survey including the original item pool, demographics, and validation measures.
- Utilize automated item selection procedures to reduce the item pool.
- Utilize an algorithm to select items from the original item pool and develop the final version of the scale.
- Use an algorithm implemented in the R package "stuart" with predefined datasets split into training and test datasets.
- Evaluate solutions against an objective function consisting of model fit criteria and composite reliability.

- Evaluate model fit using standard recommendations proposed by Hu and Bentler (1999), including χ^2 significance testing and fit indices.
- ez cutoffs. ; discussing dynamic fit indices
- Hypothesize relationships between the newly created measure and related constructs.
- Use distinct but conceptually similar instruments for validation purposes.
- Cross-validate findings using a split-sample approach?
- Evaluate solutions against an objective function consisting of model fit criteria and composite reliability.
- Validate findings using k-fold cross-validation with the dataset?
- Conduct confirmatory factor analysis (CFA) with the R package "lavaan."

3.2   3. Stage Three: External and Construct Validity Testing

- Perform bivariate correlation analyses with relevant related constructs to establish external validity.
- Investigate construct validity through multiple regression analysis, controlling for other variables.
- Discuss specific hypotheses regarding associations between the measure and related constructs.

3.3   Item Creation

A literature review was done to create items and then ten cognitive interviews were held to refine those potential items. The interviews were administered iteratively to refine the items every time a bit more. We will treat the first 25 participants like a pilot, to check for potential problems in the survey.

3.4   Rationale for Measurement Model

One decision that needs to be done by the researcher before the item selection with 'stuart', is the design of the measurement model, or how many items per factor the final scale should have. This scale will be created as parsimonious as possible, while ensuring that there is a real model fit to be estimated. So the decision in this case was of statistical nature, while with three items per factor, the model would have been just identified, four items per factor is the most parsimonious choice, where there is already a real model fit, that can be assessed in the end.

3.5   Sample

The participants are recruited on several online social media platforms. The participation is voluntary. We conducted a-priori power analysis to determine the necessary sample size for the structural equation modelling. We used the 'semPower' package in R (Moshagen & Bader, 2023) and also took a look into studies with similar goals and methods. The power analysis gave an analytical estimate for N=645, and a simulated estimate N=613, for the respective measurement model. In the literature sample sizes of N=500 up to N=1000 could be found (Algner & Lorenz, 2022; Remmert et al.,2022; Schneider et al.,2024). So the optimal sample size, we are aiming at, lies somewhere between those numbers.

Participants have to be of legal age, to be included in the study. Furthermore, attention check questions are included (three instructed response items and one seriousness check item, at the end) within the survey to assess participants' attentiveness. Participants who fail to correctly answer two out of the four attention check questions will be excluded from the analysis.

The sample for this study comprised XXX participants (M=, SD=). Within the sample, XXX% identified as female, XXX% as male, and xxx% did not identify with binary gender categories. All participants were aged 18 and above. Regarding education, all participants exhibited a [insert educational level- specifying the range or types of educational levels observed in the sample]. Among the participants, n= reported higher knowledge on items x, x, x, leading to their selection for an additional set of items as a preliminary survey for factors four and five. The study encompassed every sector within the occupational classification (Bundesagentur für Arbeit, 2020), ensuring comprehensive representation. Conducted in German, the participation in the study was entirely voluntary, with no external incentives provided. The recruitment of participants was carried out through a combination of personal and professional networks, along with outreach on various online social media platforms.

Our study sample serves as a focal point for comparison against the demographic landscape of the general public in Germany. In 2022, the mean age of the German population was 44.6 years, with 45,457,000 individuals engaged in employment. Educational backgrounds varied (XXX), and for gender distribution, the split was nearly 50/50 (41,616,473 males and 42,816,197 females) according to the Statistisches Bundesamt (source: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Bevoelkerungsstand/Tabellen/liste-zensus-geschlecht-staatsangehoerigkeit.html#651186).

3.6 Open Science Standards

This project uses the reproducibility workflow proposed by Peikert et al. (2021). Docker and renv work together to create a reproducible and portable environment. Docker captures the complete software stack, while renv focuses on managing R package dependencies and providing a clear documentation of the R package environment. This combination ensures that your analysis can be easily reproduced and shared with others in a reliable and transparent manner. The study was preregistered at Zenodo (DOI:10.5281/zenodo.11196495).

3.7 Procedure

A cross-sectional online survey is used to examine a representative sample from the general population. Participants complete the Self-perceived Data Literacy Scale along with demographic questions and additional validation measures. Survey questions of each measurement are randomized for each participant to minimize order effects and response biases. To shorten the overall length of the assessment the questions in each factor of the data literacy questionnaire are randomly selected for each participant. That way each participant only answers half of the possible items, the other half are planned missings.

3.8 Instruments

3.8.1 Measuring Data Literacy

Measure details: On Data Literacy the participants will be asked to answer 71 items. Each participant will answer 38 items of the 71, that are randomly selected. To answer the items, respondents indicate their agreement on a five-point Likert scale (1 = "strongly disagree", 2 = "somewhat disagree", 3 = "neither agree nor disagree", 4 = "somewhat agree", 5 = "strongly agree") with a "don't know" option.

3.8.2 Measuring Self-Efficacy to Information Behavior

Measure details: The SWE-IV-16 (Behm, 2018) assesses the self-efficacy beliefs of adolescents and adults in their ability to engage in information behaviour. This questionnaire measures the process model of information-related problem-solving (Brand-Gruwel et al., 2009). It consists of 16 statements addressing self-assessed abilities in searching for and evaluating information, as well as managing information searches effectively. Each statement begins with "When I search

for information on a topic or a specific question..." and respondents indicate their agreement on a five-point Likert scale (1 = "strongly disagree", 2 = "somewhat disagree", 3 = "neither agree nor disagree", 4 = "somewhat agree", 5 = "strongly agree"). Index and scoring criteria: The total scale value is computed as the arithmetic mean of the items, which may be inverted if necessary. Calculation of the total value requires valid responses to at least 12 of the 16 items.

3.8.3 Measuring Need for Cognition

Measure details: The NFC-K (Beißert et al., 2015) is a tool used to assess the NFC through four items, which represent two facets: "engagement" and "joy". The NFC-K is measured with a seven-point response scale, ranging from "strongly disagree" (1) to "strongly agree" (7), with a "neither" option in the middle. The German version of the scale is adapted from the original English scale by Cacioppo and Petty (1982) and translated by Bless et al. (1994). Index and scoring criteria: To determine an individual's NFC score, a mean value (scale value) is computed from the four raw score points of the responses. The resulting mean values range between 1 and 7.

3.8.4 Measuring Self-Perceived ICT Competency

Measure details: To assess self-perceived competence in using information and communication technology (ICT), the five general items of the ICT-SC25 (Schauffel et al., 2021) will be used. The ICT-SC25 is a scale consisting of 25 items designed to assess self-perceived competence in using information and communication technology. It is available in both German (ICT-SC25g) and English (ICT-SC25e). The scale measures general and domain-specific ICT competence, including communication, processing and storing, content generation, safe application, and problem-solving skills. Items are measured using a six-point fully-labeled Likert-type rating scale ranging from strongly disagree (1) to strongly agree (6). Index and scoring criteria: Researchers can choose to utilise either the entire scale or individual subscales based on their specific research objectives. The ICT-SC25g/e is applicable for both manifest and latent analysis. Manifest scale scores for the ICT-SC25g/e are calculated separately for each subscale by computing the unweighted mean score of the items within each subscale (Schauffel et al., 2021).

3.8.5 Measuring Openness and Conscientiousness

Measure details: The BFI-10 (Rammstedt et al., 2014) will be used to assess personality based on the five-factor model. Only the items on openness and consci-

entiousness were assessed. The items are answered on a five-point rating scale from "strongly disagree" (1) to "strongly agree" (5). Index and scoring criteria: To measure the respondent's individual traits on the five personality dimensions, the responses to the two items for each dimension are averaged. First, the negatively worded item is recoded (items 1, 3, 4, 5, and 7), then the mean value is calculated for each dimension from both the recoded and non- recoded items. The values for the five dimensions range from 1 to 5 (see Rammstedt, 2007 for reference values)

4 Analysis

Algorithm-based item selection is used to choose the most relevant items, reducing the item pool. Unlike classical approaches that consider items based on their individual merits, heuristic item selection algorithms aim to enhance the psychometric properties of a set of items within predetermined constraints (Schultze, 2017). The sample is split into training and test datasets to evaluate solutions against an objective function consisting of model fit criteria and composite reliability. In automated item selection, items can be chosen as sets that meet specific criteria. Those criteria will be defined in the objective function in 'stuart' (Schultze, 2020). I want to optimize the final model for model-fit and reliability (RMSEA, SRMR, CFI & McDonald's ω) as well as variability in the difficulty of items.

Relationships between the newly created measure and related constructs are hypothesised. Construct validity is evaluated through confirmatory factor analysis (CFA), using MLR as estimator. and correlation analyses with related constructs. Crossvalidation and measurement invariance tests are also conducted. I expect the final questionnaire to correlate moderately up to highly positive with the SWE-IV-16 (Behm, 2018), measuring peoples ability to engage in information behaviour. I expect a moderate, positive correlation of the final scale with the five general items of the ICT-SC25 (Schauffel et al., 2021). I expect a small positive correlation with the NFC-K (Beißert et al., 2015), assessing the Need for Cognition (NFC). I also expect small positive correlations with openness and conscientiousness of the BFI-10 (Rammstedt et al., 2014). The provided sample will be divided into subsets using the 'holdout' function in 'stuart'. The specified item-selection procedure will be then applied to the training dataset first. It will then be tested on the testing dataset. Validation will be conducted using the 'crossvalidate' and/or the 'kfold' function in 'stuart' to assess the invariance of the measurement models between the training and validation datasets. Invariance levels will be measured with the 'max.invariance' function. Invariance will be necessary to claim that the

scale validation has worked.

The final item selection will be determined by the highest value on the objective function in the multiple-group SEM, while ensuring 'max.invariance' between the training and validation data, as well as in a possible k-fold crossvalidation.

residuals correlates or the residuals for the adjacent constructs

5   Results

5.1   Demographic Results

On average, participants were working (M/SD). On average, participants were studying (M/SD). Descriptives and Correlations Table XX presents descriptive statistics, McDonald's ω, Cronbach's α and the correlation matrix for the respective variables.

The data literacy scale correlated XXX with the SWE-IV-16(r = , p). The data literacy scale correlated XXX with the NFC-K(r = , p). The data literacy scale correlated XXX with the the general items of the ICT-SC25(r = , p). The data literacy scale correlated XXX with the openness of the BFI-10(r = , p). The data literacy scale correlated XXX with conscientiousness of the BFI-10(r = , p). The xxx correlations (r = , p) between the newly created scale and the SWE-IV-16, NFC-K, the general items of the ICT-SC25 and openness and conscientiousness of the BFI-10,respectively, indicate the data literacy scale measures #a similar, yet distinct concept#.

5.2   Model Fit and Latent Structure in the Construction Sample

The Algorithm selected 20 of the 71 original items representing the five factors Comprehension, Evaluation, Integration, Communication and Statistics with four items each (Figure X). The final solution exhibits good model fit with Satorra-Bentler-X^2 (XXX, N = XXX) = XXX, p = XXX,CFI = XXX, TLI = XXX, SRMR = XXX, RMSEA = XXX, 90%-CIRMSEA [XXX; XXX]. Standardized loadings of the factor Comprehension ranged from 0.XX to 0.XX. For the factor Evaluation loadings ranged from 0.XX to 0.XX. For the factor Integration loadings ranged from 0.XX to 0.XX. For the factor Communication loadings ranged from 0.XX to 0.XX. For the factor Statistics loadings ranged from 0.XX to 0.XX. All factor loadings including standard errors can be found in the Supplementary Material.

5.3 Model Fit and Latent Structure in the Validation Sample

Cross-validation with the second half of the data indicated that the assumption of XXXX measurement invariance holds across the XXX subsamples: X^2(XXX) = XXX, p < XXX, CFI = 0.XXX, SRMR = 0.XXX, RMSEA = 0.XXX; X^2 = XXX, df = XX, p = 0.XXXX.

6 Discussion

- what to optimize the scale for?
- dynamic fit indices
- factors 4 and 5

  - adaptive testing/ IRT

    * dimensionality assumption and computationally intense

  - CART - tree based adaptive testing (classification trees) always binary split

    * gini index to identify the cut off
    * POMP method - for differing number of answer formats

- residuals correlates
- heterogeneity of construct

The heterogeneous nature of the construct complicates global instrument development and understanding across all participants. The measure is designed for citizens, potentially limiting discrimination at higher item difficulties or among more literate participants, a direction we aim to improve in future studies.

Noteworthy is the inherent approximate, rather than deterministic, nature of metaheuristics (Schultze & Lorenz ,2023; Blum and Roli, 2003).

References

Asch, S. (1952). *Social psychology.* Prentice Hall.

Bates, M. J. (2005). An introduction to metatheories, theories, and models. In M. J. Bates & M. N. Maack (Eds.), *Encyclopedia of library and information sciences* (2nd ed., pp. 109–121). Taylor & Francis.

Bless, H., Wänke, M., Bohner, G., Fellhauer, R., & Schwarz, N. (1994). Need for cognition: Eine skala zur erfassung von engagement und freude bei denkaufgaben [presentation and validation of a german version of the need for cognition scale]. *Zeitschrift Für Sozialpsychologie*, *25*, 147–154.

Brillouin, L. (1953). Negentropy principle of information. *Journal of Applied Physics*, *24*(9), 1152–1163.

Calzada Prado, J., & Marzal, M. Á. (2013). Incorporating data literacy into information literacy programs: Core competencies and contents. *Libri*, *63*(2), 123–134. https://doi.org/10.1515/libri-2013-0010

Carmi, E., Yates, S. J., Lockley, E., & Pawluczuk, A. (2020). Data citizenship: Rethinking data literacy in the age of disinformation, misinformation, and malinformation. *Internet Policy Review*, *9*(2). https://doi.org/10.14763/2020.2.1481

Cohen, A. R., Stotland, E., & Wolfe, D. M. (1955). An experimental investigation of need for cognition. *The Journal of Abnormal and Social Psychology*, *51*(2), 291–294. https://doi.org/10.1037/h0042761

Cui, Y., Chen, F., Lutsyk, A., Leighton, J., & Cutumisu, M. (2023). Data literacy assessments: A systematic literature review. *Assessment in Education: Principles, Policy & Practice*, *30*, 1–21. https://doi.org/10.1080/0969594X.2023.2182737

Frank, M. (2016). Data literacy - what is it and how can we make it happen? *Journal of Community Informatics*, *12*, 4–8.

Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review / Revue Internationale de Statistique*, *70*(1), 1–25. http://www.jstor.org/stable/1403713

Gould, R. (2017). Data literacy is statistical literacy. *Statistics Education Research Journal*, *16*(1), 22–25. https://doi.org/10.52041/serj.v16i1.209

Jaynes, E. T. (1957). Information theory and statistical mechanics. *Physical Review*, *106*(4), 620–630.

Kubinger, K. D., & Proyer, R. (2005). Gütekriterien. In K. Westhoff, L. J. Helfritsch, L. F. Hornke, K. D. Kubinger, F. Lang, H. Moosbrugger, A. Puschel, & G. (Testkuratorium). Reimann (Eds.), *Grundwissen für die berufsbezogene eignungsdiagnostik nach DIN 33430* (2nd ed., pp. 191–199). Pabst.

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, *50*(4), 370–396. https://doi.org/10.1037/h0054346

Murphy, G. (1947). *Personality: A biosocial approach to origins and structure.* Harper.

Payan Carreira, R., Sacau-Fontenla, A., Rebelo, H., Sebastião, L., & Pnevmatikos, D. (2022). Development and validation of a critical thinking assessment-scale short form. *Education Sciences*, *12*, 938. https:

//doi.org/10.3390/educsci12120938

Peikert, A., Van Lissa, C. J., & Brandmaier, A. M. (2021). *Reproducible research in r: A tutorial on how to do the same thing more than once.* https://doi.org/10.31234/osf.io/fwxs4

Rear, D. (2019). One size fits all? The limitations of standardised assessment in critical thinking. *Assessment & Evaluation in Higher Education, 44*, 664–675.

Sarnoff, I., & Katz, D. (1954). The motivational bases of attitude change. *The Journal of Abnormal and Social Psychology, 49*(1), 115–124. https://doi.org/10.1037/h0057453

schüller, K. (2020). *Future skills: A framework for data literacy* (Working Paper No. 53). Hochschulforum Digitalisierung. https://doi.org/10.5281/zenodo.3946067

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal, 27*(3), 379–423.

Shields, M. (2005). Information literacy, statistical literacy, data literacy. *IASSIST Quarterly, 28*(2–3), 6. https://doi.org/10.29173/iq790

Wolff, A., Gooch, D., Montaner, J. J. C., Rashid, U., & Kortuem, G. (2016). Creating an understanding of data literacy for a data-driven society. *The Journal of Community Informatics, 12*(3), 9–26. https://doi.org/10.15353/joci.v12i3.3275