ZDM – Mathematics Education
https://doi.org/10.1007/s11858-025-01665-4

ORIGINAL PAPER

Citizens' data-ing with contemporary data in their daily life

Lonneke Boels[1] · **Aline Boels**[1] · **Rosa Alberto**[1] · **Kees Hoogland**[1]

Accepted: 5 February 2025

Abstract

Citizens regularly search the Web to make informed decisions on daily life questions, like online purchases, but how they reason with the results is unknown. This reasoning involves engaging with data in ways that require statistical literacy, which is crucial for navigating contemporary data. However, many adults struggle to critically evaluate and interpret such data and make data-informed decisions. Existing literature provides limited insight into how citizens engage with web-sourced information. We investigated: *How do adults reason statistically with web-search results to answer daily life questions?* In this case study, we observed and interviewed three vocationally educated adults searching for products or mortgages. Unlike data producers, consumers handle pre-existing, often ambiguous data with unclear populations and no single dataset. Participants encountered unstructured (weblinks) and structured data (prices). We analysed their reasoning and the process of preparing data, which is part of data-ing. Key data-ing actions included judging relevance and trustworthiness of the data and using proxy variables when relevant data were missing (e.g., price for product quality). Participants' statistical reasoning was mainly informal. For example, they reasoned about association but did not calculate a measure of it, nor assess underlying distributions. This study theoretically contributes to understanding data-ing and why contemporary data may necessitate updating the investigative cycle. As current education focuses mainly on producers' tasks, we advocate including consumers' tasks by using authentic contexts (e.g., music, environment, deferred payment) to promote data exploration, informal statistical reasoning, and critical web-search skills—including selecting and filtering information, identifying bias, and evaluating sources.

Keywords Informal statistical reasoning · Consumer task · Producer task · Internet search · Vocational education · Statistical literacy

1 Introduction

The Web and technological advancements [e.g., search engines, Artificial Intelligence (AI)] have changed the data and information people encounter. To make informed decisions in their daily life, citizens rarely produce data themselves or conduct statistical analyses. Instead, they search the Web. The multivariate data they find are often pre-processed, may have been created for other purposes, or may not have been purposefully produced (Gould, 2021). Thus, citizens can be considered *consumers* or users of data, rather than data producers (Gal, 2000; Gigerenzer, 2022; Gould, 2017).

However, a substantial portion of the population struggle when exploring, selecting, and interpreting contemporary data, information, and statistics (Boels et al., 2022). In the Netherlands, 15.5% of adults (aged 16–65) scored at level 1 or below on numeracy in the last Programme for the International Assessment of Adult Competencies (PIAAC) (OECD, 2024). This suggests adults likely have difficulties to "understand a broad range of mathematical [including statistical] information that may be complex, abstract, or found in unfamiliar contexts" and struggle to "interpret and perform basic analyses of data and statistics in texts, tables, and graphs" (OECD, 2013, p. 2).

✉ Lonneke Boels
Lonneke.boels@hu.nl

Aline Boels
alineboels@live.nl

Rosa Alberto
rosa.alberto@hu.nl

Kees Hoogland
kees.hoogland@hu.nl

[1] HU University of Applied Sciences Utrecht, Utrecht, The Netherlands

Published online: 10 March 2025

Statistical literacy—the ability "to interpret, critically evaluate, and communicate about statistical information and messages" (Gal, 2002, p. 1)—is vital for several daily life decisions (e.g., during high inflation). To interpret contemporary data, citizens must integrate numerical, mathematical, and statistical information (Gal & Geiger, 2022). It requires them to consider the data context and critically evaluate what the data are, where they came from, and how they can—or cannot—contribute to inform their daily decisions.

Digitisation and technological advancement have changed the competence citizens need over the past 20 years (Gal & Geiger, 2022; Hoogland & Stelwagen, 2021). In the twenty-first century, data interpretation increasingly requires digital competence and interpreting numbers and texts (Hoogland & Díez-Palomar, 2022; Ridgway et al., 2011). While the data may not always be new, the way they are delivered and made accessible has changed considerably (e.g., digitally, in a database, or through weblinks) alongside an exponential increase in available data (e.g., real-time stock data in 'webshops' (online retailers) versus annual inventory).

The existing body of knowledge offers limited or outdated examples of contemporary data situations people encounter in their professions (Bakker, 2014; Kanes, 1996) and daily lives (Gal & Geiger, 2022) or how people use statistics and numeracy to interpret and reason with contemporary data (e.g., Geiger et al., 2015). Few empirical studies exist (Gal, 2024a).

This article explores how citizens engage with contemporary data and reason statistically as part of their statistical literacy. We examine how they make sense of search results including consumer reviews and statistics (e.g., mean product review score). For instance, citizens seeking to purchase goods may need to critically assess search results, including texts, images, and numerical data from reviews, magazines, adverts, consumer unions, and online stores.

This study explores citizens' statistical reasoning in daily life by examining how they interpret contemporary data in life contexts to answer their own questions through Web searches. The research question is: *How do adults reason statistically with web-search results to answer daily life questions?* The focus is on adults up to and including vocational education—specifically those with little statistical training. This part of the population is underrepresented in research (Gal et al., 2020; Gal, 2024a).

2 Theoretical background

In this theoretical background section, we discuss more in detail what data and data producers are, what statistical reasoning, informal reasoning, and web-search results entail, and what data-ing is. As underlined by Gal (2024b), consumer tasks, such as those in our study, differ from the producer tasks in statistics curricula.

2.1 Data

Data are traditionally defined as a collection of recorded observations. Data and information are often used as synonyms, as seen in Wise's (2020) definition: "Data can be broadly defined as information collected or generated from the world from which inferences about various phenomena can be made" (p. 165). Some fields, like economics, distinguish between *data*—rooted in "differences in physical states-of-the-world"—and the *information* extracted from them through interpretation (Boisot & Canals, 2004, p. 46). When direct observation is not possible, *proxy variables* are often used—substitutes for variables that are hard or impossible to measure directly. For example, in social science, time spent on a task is used as a proxy for students' active participation (Kim et al., 2016); a mercury thermometer uses mercury expansion as a proxy for temperature.

Traditionally, data are produced by researchers, national statistics institutions, or governments. The increasing availability of data on the Web now allows data scientists to answer questions by analysing data from disparate sources which are produced by others for a different application (Gould, 2021). These data may be obtained through *data harvesting* (Fry & Makar, 2021; IDSSP, 2019) which includes Web scraping of unstructured data and structured datasets (Cafarella et al., 2011). Structured data are highly organised, typically in tables (e.g., rows and columns). Data scientists employ techniques such as classification, regression, clustering, and association—often using AI tools—for analysing such data. They also use data moves and pay attention to communicating results (Fry & Makar, 2021; IDSSP, 2019).

Data moves are actions that prepare data for analysis such as selecting, filtering, calculating new attributes, grouping, summarising (e.g., in tables, graphs, measures of central tendency), and changing a dataset's contents, structure, or values (Erickson et al., 2019). In data science, "data are *chosen*, *transformed*, [emphasis in original] and even *created* by humans, and … are dynamic and changeable" (Strohmayer & Muller, 2023, p. 39) which impacts analysis and interpretation (Gould, 2017). Important for data analysis are: the context of the data (e.g., who produced it, how, and why) (Garfield et al., 2008), the type of data (images, texts, numbers) (Garfield et al., 2008), whether data are aggregated (e.g., in tables and graphs), and whether the data are or could have been altered (e.g., through data moves; Erickson et al., 2019).

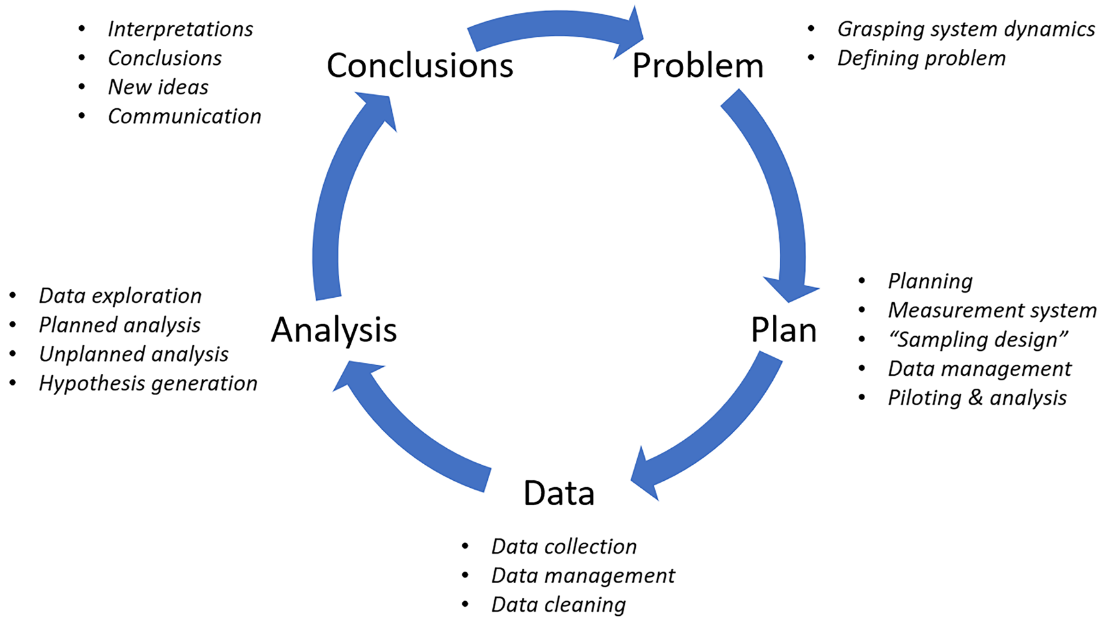

Fig. 1 Investigative cycle (Wild & Pfannkuch, 1999)

2.2 Data producers

The literature distinguishes between data producers and data consumers (e.g., Engel & Ridgway, 2022), aligning with Gal's (2000) description of statistical literacy as "people's ability to interpret and critically evaluate statistical information and data-based arguments [… and] emphasiz[ing…] sense-making … more than formal statistical knowledge, assuming that most adults are consumers rather than producers of statistical information" (p. 135). Gal (2024b) contrasts adults' consumer tasks with producer tasks learned at school. His earlier work related producer tasks to "reporting contexts [that] emerge when learners are 'data producers' and take part in all phases of a data-based study" (p. 137). However, adults "have to make sense of and possibly react to messages that contain statistical elements." Such messages stem from "listening (reading contexts) [that] emerge when learners are 'data consumers'" (p. 137). In addition, Engel and Ridgway (2022) advocate "rethinking the purpose and nature of statistics education" as "many introductory statistics courses fail to prepare students to understand and critically analyse empirical evidence in the public domain" (p. 17), a consumer task.

We illustrate this focus on producer tasks with an example of young people's happiness. We follow the investigative cycle (Wild & Pfannkuch, 1999, Fig. 1). Suppose we hypothesise that boys are less happy than girls (*Problem*). There are many ways to operationalise happiness (*Plan*), such as asking about life satisfaction, cheerfulness, positive or negative emotions, feeling happy, or by measuring heart rate, cortisol levels, blood pressure, etc. In addition, decisions are needed on scales (words, numbers, emoticons), scale lengths, and single or repeated measurements. Data from other sources (parents, teachers, and researchers' observations) could be included. It requires decisions on age, power analysis (expected effect size), sample size, and methods. Sampling methods (e.g., online surveys) may exclude certain groups, like those from low-income families. Next, decisions need to be made about *Data* storage and missing or incorrectly entered data. *Analysis* involves selecting and interpreting appropriate graphs that show data distribution, variation, and outliers and deciding if further analysis is needed. It may involve hypotheses testing or calculating means, standard deviations, or associations. Finally, conclusions are drawn from these analyses, such as whether there is a significant difference (*Conclusions*). Conclusions depend not only on analyses but also on initial choices (*Plan*) showing the impact of the data producer's decisions. These processes, decisions, and ambiguities are typically not visible to data consumers.

2.3 Data-ing

Data-ing is a new term for the process of engaging or reasoning with data (SRTL13, 2022). We focus on the part of this process that involves actions, decisions, and judgements to obtain and prepare data for analysis. It includes deciding how to operationalise the phenomenon, judging the relevance of variables, deciding on suitable proxy variables when direct measurement is not possible, and deciding on sample size or how much data is needed (*Plan*, Fig. 1). It also includes data moves (Erickson et al., 2019) such as selecting multiple and diverse data, filtering data, calculating new attributes (*Data Management)*, grouping, and summarising (e.g., in tables, graphs, measures of central tendency) to prepare for analysis. The latter part aligns with *Data exploration* (Fig. 1) but in our view of data-ing, this fits more naturally into the *Data* step of the cycle. In our study with citizens, the data-ing part that includes all preparatory actions for analysis, is most relevant. Although reasoning about data occurs throughout all steps of the investigative cycle (Biehler et al., 2018) when analysing citizens' data-ing, we focus on the initial steps before analysis.

2.4 Statistical reasoning

The focus on data production is latent in the different views on statistical reasoning. Statistical reasoning involves interpreting data, visual representations, and statistical summaries (delMas, 2004) and using statistical concepts to make sense of statistical information (Garfield, 2003). It requires understanding statistical processes and making connections between concepts (Ben-Zvi & Garfield, 2004). Lovett (2013) adds that statistical reasoning involves tools to summarise, predict, and conclude from data. delMas et al. (2006) developed the Comprehensive Assessment of Outcomes in Statistics to measure students' statistical literacy and reasoning about descriptive statistics, bivariate data, probability, and statistical inference. Garfield's (2003) Statistical Reasoning Assessment covers reasoning about data, representations, statistical measures, uncertainty, samples, and association. Statistical reasoning includes understanding uncertainty and significance (Gal & Geiger, 2022), bias (delMas, 2004), and data as a combination of signal and noise (Ben-Zvi & Garfield, 2004; Engel et al., 2008). These views emphasise the use of formal statistical tools, like hypothesis testing and calculating and comparing statistical measures of centre and variation, that are central to data producers.

In addition, although there are similarities, mathematical and statistical reasoning differ (delMas, 2004; Garfield, 2003). For example, in mathematics, context can be omitted but in statistics, it cannot as the context informs procedures and interpretation. In contrast to mathematics, data are often messy. In statistics, there is no single answer to an investigative question and uncertainty is inherently part of it (Garfield, 2003). Therefore, unlike mathematical procedures, context must always be considered.

We use the term *informal statistical reasoning*. The word informal emphasises that formal statistical procedures (e.g., calculating association or standard deviation) are not used. For example, when participants reason about the representativeness of review results (in a webshop) they could draw on personal knowledge (a review can be written out of anger, see supplementary information for more examples). A large body of research in statistics education is focused on *informal inferential reasoning* (IIR). We suggest IIR to be a subset of informal statistical reasoning. Inferential reasoning relates to parameter estimation, comparing samples, and generalisation from sample to population (Zieffler et al., 2008). Zieffler et al. define informal knowledge as adults' understanding from daily life and less formalised knowledge from previous education. Rufiana et al. (2018) used a narrower interpretation of the word 'informal' in 'informal statistical reasoning' referring specifically to students' use of everyday language in an introductory statistics course.

2.5 Web-search results

This article focuses on tasks that require consumers to search the Web to answer their own investigative questions. Web results include image-based (e.g., pictures of products), semi-structured (e.g., scraped from webpages with reviews), unstructured (weblinks), or repurposed data. These data do not fit the traditional definition of data as a collection of recorded observations (e.g., Wise, 2020). As a preliminary position, we consider search results to be contemporary data.

People believe that Web results are ordered by relevance, whereas it is highly influenced by search engine optimisation (Lewandoski & Schultheiß, 2023). Web results are often sponsored (Pan et al., 2007). In addition, personalisation through cookie settings and past search history can influence which results are shown and in which order (Urman & Makhortykh, 2023). Furthermore, due to randomisation (Urman & Makhortykh, 2023) two searches done by the same person on the same laptop with the same search words within two minutes may result in different links or a different order of these links. This may have a considerable effect on user behaviour (Lewandoski & Schultheiß, 2023; Pan et al., 2007).

Many webpages do not contain a database, only unstructured data. However, structured data can be found on the Web (Cafarella et al., 2011) for example, in webshops. Although structured data on the web, including those in webshops, share similarities with traditional database systems, they have some different characteristics: they are "embedded in textual webpages and [they] must be extracted prior to use; there is no centralized data design as there is in a traditional database; and, unlike traditional databases that focus on a single domain, it covers everything" (p. 72). For websites that contain structured databases, filtering actions lead to multiple choices for a product, in a specific order—where it is unknown to the user how this order was determined and influenced by an (AI) algorithm like a recommender system (Necula & Păvăloaia, 2023).

2.6 Web-search results, data-ing, and statistical reasoning

To the best of our knowledge, research on adults' data-ing and statistical reasoning with contemporary data is yet to be developed. However, related fields are relevant to citizens' statistical reasoning. Firstly, civic statistics has similarities with web-search results. Civic statistics focuses on data-based evidence relevant to society. Engel and Ridgway (2022) wrote about civic statistics: "Statistical information about society is often quite complex. Data are usually multivariate; aggregated data and indicator systems are common; variables interact; data may be time critical", and data are often buried in texts (p. 22). Furthermore, "it is unusual for a single data set or a single analysis to be sufficient to answer a question in the arena of Civic Statistics" (p. 24). This is also often the case for the questions to which citizens seek answers on the Web. Moreover, decisions made about how a certain phenomenon is measured are often opaque and data may be aggregated. Such data are messy and their interpretations may vary (Ben-Zvi & Garfield, 2004).

Secondly, a relevant field is data and statistics in (social) media. Gal and Geiger (2022) analysed media items related to the COVID-19 pandemic, which included Statistical and Mathematical Products (StaMPs). These StaMPs place interpretive and evaluative demands on readers or viewers. StaMPs included visual representations of data and statistics, often used to make texts more appealing. Important themes in media items were data quality and strength of evidence, suggesting that the ability to question both is crucial for citizens. The dynamic nature of pandemic-related statistics also highlights the importance of understanding the tentative nature of scientific findings.

Thirdly, people's search behaviour on the Web is investigated either through eye-tracking or log files. Eye-tracking research is typically small-scale, lab-based, and focuses on individuals, while log-file research aggregates data over multiple participants (Urman & Makhortykh, 2023). Three types of search tasks are described: informational (searching for information), navigational (navigating to a specific webpage), and transactional (performing a transaction) (Strzelecki, 2020). Eye-tracking studies mainly used Google and investigated search results without sponsored links, finding four types of search strategies for these tasks: breadth-first (checking multiple results and then opening the most promising), depth-first (start at the top and decide to open it or not), only selecting few top results, and others. Most studies included university students only, a major limitation in this body of literature. Log-file analysis showed that in Switzerland and Germany, almost all users clicked on the first page of search results, with Google being the most used search engine. Top results received disproportionately more clicks, with about half on the top first result (Urman & Makhortykh, 2023). Log files do not describe people's reasoning with the results or how they make sense of the data. The present study addresses this gap in the literature for adults.

We focus on the data-ing part that includes all the preparatory steps for analysis in the context of citizens' engagement with contemporary data. This includes what people do to get the data they need and using proxy variables. It includes thinking about aspects of data that are important for analysis and data moves (Erickson et al., 2019), like selecting, filtering, sorting, and ungrouping data (e.g., from a mean rating to individual ratings).

3 Method

3.1 Context and participants

Three adults, each 36 years old with a vocational background, participated in this observational multiple case study (see Table 1). Due to the exploratory nature of this study, we used a convenience sample recruited through our professional and personal network. All participants gave their consent and each received a 20-euro gift card. Data collected during the introduction of the interview (e.g., on their job or daily life numeracy experience) or at the end (when participants talked about other Web searches without applying them) were coded when relevant.

Table 1 Overview of participants' professions and questions

Name	Interview duration (h:m)	Profession	Question 1	Question 2
Bo	0:31	Host at an institution for people with intellectual and physical disabilities	What cleaning products should I buy for the company I work for?	Recipe sweet and sour chicken Recipe marble cake
Sem	1:04	Phlebotomist (someone who collects blood samples)	What vacuum cleaner should I buy for my family?	
Isa	0:59	Civics teacher	How much of a mortgage can I get? Follow-up question: Can I get a mortgage with a study debt?	What kind of house could I buy on my income?

3.2 Tasks and procedure

All interviews followed an observation and interview protocol with clarifying questions like: What are you looking at, why did you click on that, what are you paying attention to, what do you notice, what information are you looking for? At the start of the session, we explained our focus on how people search the Web in situations where money, time, numbers, and graphs may play a role. We collected demographic data such as participant's job and age. Next, participants formulated a daily life question they wanted to investigate (see Table 1). Some examples were provided, including:

- What is the best hand blender to buy?
- I'm renting now and considering buying a house. What will it cost me per month?
- How do I remove weeds, moss, or algae from my garden or balcony in the most environmentally friendly way possible?

Participants searched until they answered their questions, then continued with a second question if time allowed.

3.3 Data collection

The face-to-face interviews were audio-taped with a Zoom H5 portable audio-recorder[1] and recorded with a video recorder, or recorded in Microsoft Teams. The recordings were transcribed with the automatic transcription application Amberscript[2] and manually corrected. The data consist of transcripts of interviews and screen recordings or videos of participants interacting with the search engine and websites. Researchers interviewed participants in a library or at their homes. An HP-ProBook laptop of the researcher was used for searching the Web with Google.

3.4 Data analysis

We analysed what data participants encountered, what search terms they used, and how they reasoned with the data found during their Web search with Google and subsequently visited pages. When analysing their reasoning, we alternated between emergent interpretations and existing explanations using a pragmatic iterative approach (Tracy, 2013) as follows. For the emergent interpretations, we used open coding (inductive approach, Twining et al., 2017). By grouping closely related open codes into broader codes, subcategories were obtained (Corbin & Strauss, 1990; sometimes also called axial codes) that are interpretive (cf. Tracy, 2013). For example, the open codes *filtering of data* and *sorting of data* were merged into the subcategory *data moves* (cf. Erickson et al., 2019). All subcategories that emerged during our analysis can be found in the online Supplementary Information, with an example for each subcategory.

For the existing explanations, we used the investigative cycle (Wild & Pfannkuch, 1999), statistical reasoning (Ben-Zvi & Garfield, 2004; delMas, 2004; delMas et al., 2006; Garfield, 2003; Lovett, 2013), and notions of informal (inferential) reasoning (cf. Zieffler et al., 2008). Three final categories emerged: *choosing data*, *analysing the data,* and *drawing conclusions from the data*, see the online Supplementary Information. These categories have similarities with the investigative cycle (Wild & Pfannkuch, 1999). In addition, the subcategories in *choosing data* fit our focus of data-ing (i.e., actions, decisions, and judgements that prepare for analysis) and are elaborated in the results section. Furthermore, *analysing and drawing conclusions from data* are related to statistical reasoning.

Coding was done by one of the authors (first coder; not involved in the observations), then discussed with another author (second coder), and eventually within the team. After coding all three transcripts, the first coder went back to check for subcategories that emerged in later parts of transcripts (cf. Corbin & Strauss, 1990). We discussed the subcategories for specific parts of the transcripts until all researchers reached an agreement. The videos were annotated for key events including where participants clicked, what data they encountered, and what search terms they used. These annotations were added to the transcripts where relevant.

4 Results

Three participants searched the Web to answer daily life questions: Sem for a vacuum cleaner, Bo for a cleaning product and recipe, and Isa to assess house affordability. We explored: *How do adults reason statistically with web-search results to answer daily life questions?* We used transcripts of participants' verbalisations during Google searches and visited pages. We focused on (a) types of data encountered, (b) participants' data-ing, specifically all the actions, decisions, and judgements that prepare for analysis, and (c) their statistical reasoning during data analysis and drawing conclusions.

4.1 Data: types of web-search results

In Google, participants searched with queries like *vacuum cleaner family* (Sem), *kitchen cleaner* (Bo), *marble cake recipe* (Bo), *calculate mortgage* (Isa), and *how to calculate percentages with calculator* (Isa). Figure 2 shows the

[1] https://www.thomann.de/nl/zoom_h5.htm.

[2] https://www.amberscript.com/en/.

results when entering the query *cleaning products*. When participants searched for products, Google presented first links as images with text (e.g., product prices) followed by textual links. These unstructured data (Cafarella et al., 2011) may vary in time and depend on randomisation and cookie settings (Urman & Makhortykh, 2023). Such data also lack a consistent format. Therefore, these data are considered messy and difficult to interpret or analyse. As many first links are sponsored, users must critically evaluate them for relevance, bias, and trustworthiness (see Sect. 4.2.1).

Next, participants accessed different websites, including webshops (Bo, Sem), homes website (Isa), and YouTube (Bo). These sites contain structured data on prices, reviews, and meta-data (e.g., author, views) (cf. Cafarella et al., 2011) automatically extracted from databases (Arasu & Garcia-Molina, 2003). Unstructured data were found on websites with, for example, mortgage calculators (visited by Isa).

Notably, Web data often lack clear producers (e.g., researchers), samples, or populations. In addition, there was no single dataset with predefined variables. What data moves (Erickson et al., 2019) had been applied was unclear. Data of products on top of the Google search page stemmed from many different webshops (Fig. 2). Data in webshops are updated irregularly (e.g., reviews posted at any time, products added or removed without notice). The data are often from multiple sources (e.g., consumers, other websites) with reviews sometimes scraped from the manufacturer's website (Fig. 3).

4.2 Data-ing

We focused on the data-ing part that involves actions, decisions, and judgements to obtain and prepare data for analysis. In the investigative cycle (Wild & Pfannkuch, 1999) these steps include defining the problem (*Problem*), determining how to operationalise variables and what data are needed (*Plan*), and searching for or producing these (*Data*). We captured these steps with the category *choosing data*. Although the first three steps of the investigative cycle (Wild & Pfannkuch, 1999) seem to align, the focus on *choosing data* (Strohmayer & Muller, 2023) and consumers' tasks reveals a different practice than originally intended. Within this category several subcategories emerged from participants' reasoning that are relevant to data-ing: evaluating web-search results, using data moves (Erickson et al., 2019), and using proxy variables.

4.2.1 Data-ing: evaluating web search results

Evaluating and judging the relevance of Web search results is an important aspect of citizens' critical use of data. That includes deciding where to click.

Bo Most of the time I look at the first one [clicks on the first recipe, goes to leukerecepten.nl, scrolls through the webpage].

In Google, participants often clicked on the first or first relevant result. Participants also used personal knowledge and personal experiences with webshops, products, or brands. For example, Bo looked at Google images (Bo said to compare prices and cleaning products' effectiveness) and then clicked on a sponsored link from a familiar brand. Explained by Sem:

Sem I am always inclined to check Bol.com. This is purely because I know that if I order something there, nine times out of ten I have it the next day.

Sem chose to include the name of a well-known webshop in the Google search query and clicked on the sponsored link. Isa also clicked on the first link Google provided (sponsored). However, when the visited site asked to fill in the specific mortgage amount they were considering, Isa realised they did not know. Therefore, they returned to the Google search results and selected a link they recognised from the radio: *Calculate Your Maximum Mortgage*.

Sometimes participants deliberately chose data sources (e.g., a bank, website). In Google, they often showed an uncritical approach by clicking on the first or sponsored results. Bo ignored search results for kitchen cleaners, due to a preference for a specific brand. This shows that data are not always critically chosen. Consistent with the literature on online search behaviour, participants clicked disproportionately on the top results (Strzelecki, 2020; Urman & Makhortykh, 2023), underscoring the need to address this in education.

All participants explained which results they could use, for example:

Bo I also usually look at how long … ago it was posted [e.g., a video on how to bake a cake]. If it's too much from 10 years ago, I would never [grab …] that first. Look at most recent.

In the above excerpt, Bo judged whether data provided by YouTube were relevant, based on publication date. Another example of critically assessing the data is provided by Sem.

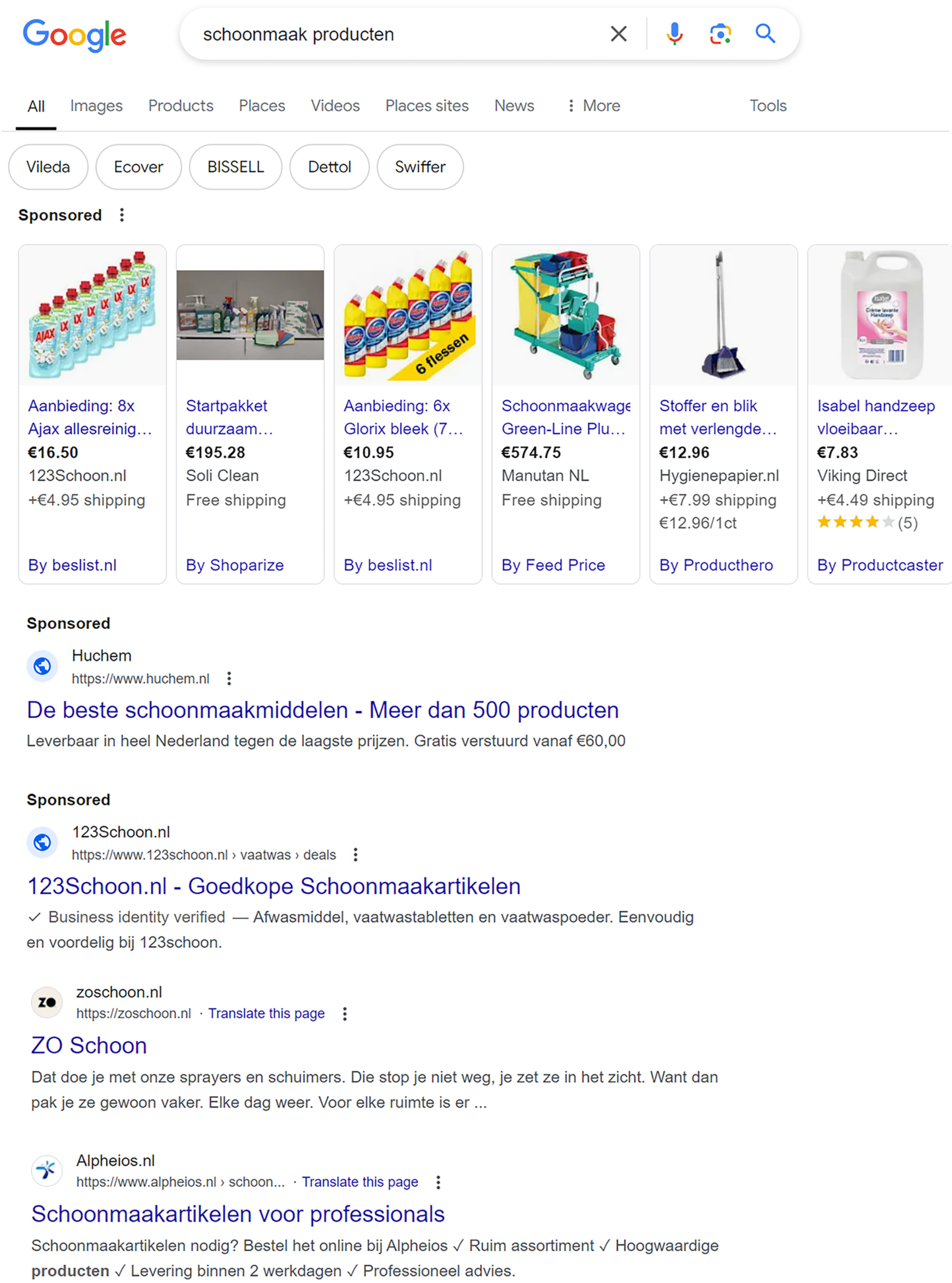

Fig. 2 Google search results for *cleaning products*. Note: Sponsored images are at the top, followed by sponsored links (retrieved with Google, October 2024)

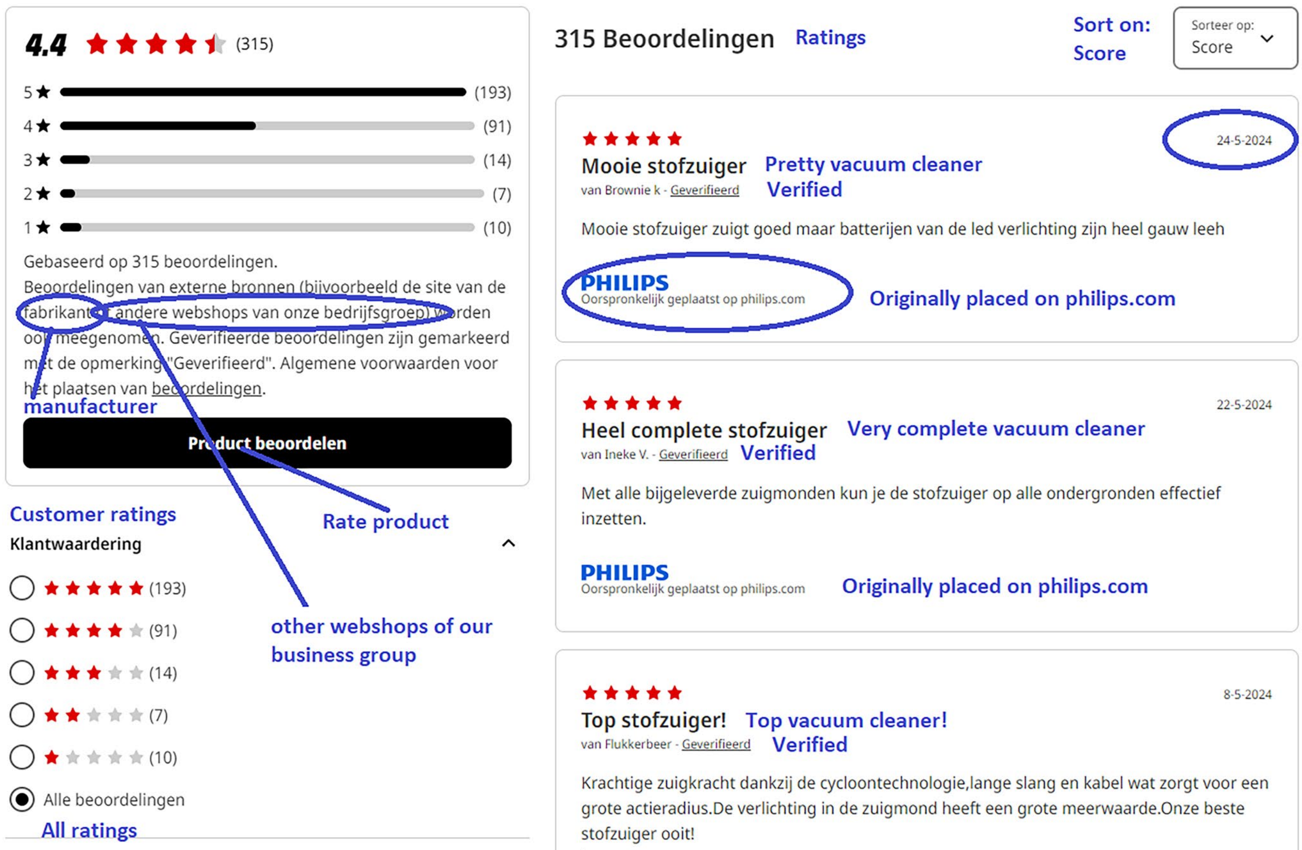

Fig. 3 Reviews are added on an irregular basis (see dates, e.g., top right) and in part scraped from other websites (e.g., website manufacturer). Source: Webshop M

Sem What I find strange is the six-meter cord [product description]. I just saw it listed as nine meters [results overview]. … I would think the range was based on the cord length, so I feel a bit cheated reading now: six.

Sem noticed that the webshop provided two measures of the vacuum cleaner's reach: first 9 m (action radius) and 6 m (cord length) which raised concerns about the website's trustworthiness. Participants discussed which websites, search results, or products they trusted and why. While they were uncritical when using Google, they became more critical when visiting subsequent websites, and evaluating data for trustworthiness, relevance, and usefulness. Wild and Pfannkuch (1999) consider critical thinking necessary at all stages of the investigative cycle, aligning with broader discussions in data science on the importance of understanding data sources (Gould, 2017; Strohmayer & Muller, 2023).

4.2.2 Data-ing: using data moves

Webshops and homes websites (and, to a lesser extent, YouTube) offer many *data moves* (Erickson et al., 2019), especially selecting, filtering, and sorting (e.g., by brand, noise level, battery life, price range, and customer reviews (1 to 5 stars) or on homes websites by house type, living area, or having a garden). Participants discussed why they filtered data in webshops. For example:

Sem [Selects from the menu *bagless vacuum cleaners*] I find that more practical and I think it will result in less need to buy bags.

Sem extensively filtered vacuum cleaners, considered and selected several criteria from the menu including *in-stock items* (reason: they needed immediate replacement of the vacuum cleaner), kept *all floor types* (as their home has a mix of carpet, tile, and wood), and wished they could filter by *cord length* (to avoid switching plugs). They note that aesthetics are unimportant for a vacuum cleaner as it is subject to rough handling. Sem filtered vacuum cleaners on *three or more stars*, explaining that a five-star rating would result in expensive vacuum cleaners, while others offer better value for money. Sem also once sorted results in *ascending price order*.

Isa checked the relevance and filtered data less often and less explicitly than Bo and Sem. Isa initially did not filter on the homes website. The first home displayed on Funda is €514,000 which Isa considers to be quite high. The second home was so exorbitant (€1,189,000) that they struggled to pronounce the amount. Isa was prompted to filter so as to adjust the price range within the mortgage amount of €136,000.

Isa Yes and then let's take a look at the 136 [thousand]. Let's take the cheapest option for a moment [clicks on *price*; sets 'from' at €50,000]. So, we'll start at 50,000 and then the maximum would be this, I think [clicks on *price*; sets 'to' at €125,000].

This adjustment led them to a garage. They then increased the maximum limit to €150,000 but the resulting home was not an improvement over their current rental situation.

Participants used the data moves (Erickson et al., 2019) provided by the websites and search engines, demonstrating varying skill levels. The selecting and filtering produced multiple options in a specific order, though participants did not discuss or seem aware of how this order was determined or influenced by AI algorithms (e.g., a recommender system). This sometimes led to surprises, such as finding expensive houses. These results show that it is essential that users learn to critically evaluate the results (e.g., Gal, 2000; Gould, 2021), filter them for relevance (e.g., Gal & Geiger, 2022), and carefully interpret them to make evidence-informed decisions (e.g., Engel & Ridgway, 2022; Gal, 2000; Gal & Geiger, 2022).

4.2.3 Data-ing: using proxy variables

Participants sometimes used proxy variables: substitutes for variables that are not available on the sites they are visiting (cf. Kim et al., 2016).

Sem So yes, 2200 Watt [vacuum cleaner at home …] then I think, it [description of vacuum cleaners in search results] says 700 or 900 Watt … . So, this could mean that these vacuum cleaners are actually very weak [in suction].

Sem used electric power (Watt) as a proxy for a vacuum cleaner suction power as suction power was not available.

Discussing what videos are trustworthy, Bo stated:

Bo And I also look at the views. When so many people have watched it, I usually consider that as well.

Bo uses the number of views (video) as a proxy for trustworthiness. Participants used proxy variables when other data were unavailable and alternated between what they wanted to analyse (e.g., suction power) and what variable was available as a proxy (e.g., electric power). Other examples of proxy variables used by participants were price for quality of the product, tone of voice for quality of the review, and year of publication (of a review, video) for relevance.

4.3 Statistical reasoning during data analysis and drawing conclusions

In the investigative cycle (Wild & Pfannkuch, 1999) formal techniques are used for data analyses like making tables, calculating statistical measures, and creating graphs to reveal patterns. Most data found by our participants cannot easily be analysed with such techniques. The Web scraping tools used by data scientists—that could allow formal analysis—are out of reach for most citizens. As we saw in the previous and elaborate in the following sections, our participants handled, reasoned with, and analysed data differently.

Many aspects of formal statistical reasoning associated with producers' tasks (e.g., Ben-Zvi & Garfield, 2004; Garfield, 2003) were absent because participants were engaged in consumers' tasks. Moreover, participants' reasoning remained informal. In total, only four categories from formal statistical reasoning (e.g., delMas et al., 2006; Garfield, 2003) emerged: *reasoning about statistical measures, association, samples,* and visual *representations of data* (only tables).

4.3.1 Reasoning about statistical measures

In the following excerpts, Sem and Bo state that star's ratings are a mean:

Sem You can choose a rating of one star, two, three, four, or five. … I opt for three-star and above. [The rating of a vacuum cleaner] Is actually a mean.

Bo I usually judge it [reviews] … on … [the] mean and then I continue with that.

Sem explained that a rating of three stars does not imply that all reviewers gave three stars. As stated by both, the rating is a mean, which is indeed true (Fig. 3). Participants reasoned with statistical measures like mean, range, minimum, and variation. Bo and Sem reasoned with these concepts several times while Isa did so twice (cheapest, maximum).

Although participants used statistical measures such as the mean and range when evaluating product reviews, their reasoning did not align with what is typically assessed in statistical reasoning tests (e.g., delMas et al., 2006). Rather than engaging in formal procedures such as hypothesis testing or graph interpretation, participants took an informal approach. Similar to informal inferential reasoning (Zieffler et al., 2008) we consider this to be part of informal statistical reasoning.

Sometimes misinterpretations occurred:

Bo Right that's here, I'm looking at that too [Trustpilot review score 4/5 stars in an image, *TrustScore 3.9, 6,353 reviews* in text].

Researcher 1: … And what does that tell you, that?

Bo That it [the product they are looking at] is good stuff though.

Bo mistakenly believed that the Trustpilot score applied to the product, rather than to the webshop as a whole.

4.3.2 Reasoning about samples

Some reasoning about samples also occurred, especially when it came to the reviews, as the following excerpts illustrate.

Sem I take into account that I select three or more [stars], because if someone has had a wrong delivery or received the wrong colour, that person could be upset and might choose a one-star rating. Therefore, there should always be nuance in it.

Bo Sometimes yes and sometimes not [reviews are relevant], because you don't know how honest people are. … You can also write something out of anger and then it's, then you don't know what you [Bo can] get out of it.

Sem If I were to buy a car now, from a garage, I definitely read the reviews. Whether there is a lot of saying of well, I bought that, and nine out of 10 people all had complications with the car within three months. I look at that too.

Sem said they use reviews for expensive purchases, but not for a vacuum cleaner. Both Sem and Bo seemed aware that reviews can be biased and influenced by other factors (e.g., emotions) which indicates that they thought about the representativeness of a sample. Sem also used "a lot" which could hint at an assumption that more reviews points to greater validity or reliability. Participants' reasoning about samples was very limited. For example, what is left unspoken of the representativeness of the sample, is that reviews are only given by a part of the buyers. Therefore, it is a (non-random) sample (of a population of buyers). Gould (2017) already argued that interpreting data from non-random samples is an important part of statistical literacy for students and adults. Such a consumer task deviates from what is required for producer tasks (e.g., deciding on sample size through power analysis, choosing sampling methods that make the data more representative, making inferences about unknown populations) (e.g., Bargagliotti et al., 2020; Biehler et al., 2018; delMas et al., 2006; Wild & Pfannkuch, 1999).

4.3.3 Reasoning about association and comparing data

Important aspects of reasoning during Web searches included comparisons and associations of the provided data (e.g., product features in a webshop). The following excerpt shows how Sem compared data of the first and second vacuum cleaner in the list.

Sem My first option was 85 euros, second option 109. …I have 27 items [in this selection] so I'm annoyed a tiny bit because that means that [with] article five, I'm already at 190 euros [… which is] a difference of 105 euros from the first article.

Sem So, am I a little bit brand sensitive when I do not know the brand at all … . I see a Koenic here. I've never heard of it before. That's the first option and I chose ascending [price], so I immediately realise that the second option is a Samsung. It costs 109 euros, so there's already a price difference of 24 euros. … if I choose a Samsung, I know I will always be able to find the parts or accessories everywhere.

Sem related these results to price, price differences, count of relevant items, and their implicit, maximum budget. Sem connected qualitative (brands) and quantitative (prices) attributes which can be seen as reasoning about an association. Sem considered future repairs but did not check spare part availability, hence, relied on personal knowledge. We regard this attention to future use (e.g., spare parts) as part of data-ing.

Reasoning about associations is part of statistical reasoning (e.g., delMas et al., 2006). This reasoning also includes being able to distinguish association from causation (Garfield et al., 2008). However, the informal way in which participants reasoned about associations does not fit well with descriptions of associations such as being able to interpret a two-way table or a scatterplot (e.g., Garfield,

2003) as neither were provided. In addition, reasoning about an association between quantitative and qualitative data is not part of the Dutch secondary and vocational education curriculum.

In the following excerpt, Isa interprets the data found for the maximum mortgage.

Isa And the gross monthly expenses [of a mortgage of €136,149] would be €650. … So, about the same … what I pay in rent.

Isa compares monthly mortgage expenses with their own rent. This excerpt is an example of how participants compared data. Another website offered Isa an alternative mortgage of €165,418. Checking data on another website is also part of being critical about data on the Web (see Sect. 4.2.1).

4.3.4 Reasoning about representations

Webshop M provides users with a table with three items selected for comparison with most settings predetermined (e.g., what and how many variables per product).

Sem Well, then I see three columns. … It's a table where the specifications are listed on the left, requirements … . And then, for all three it says what is included or what is the range in this case. Whatever. Well, then I look at the three comparisons, is not very handy, because you already lose track of which one it was.

Sem reasoned that this table needs modification because it does not provide a good overview. Visual representations of data were rarely used, except by Sem who used a table. Websites generally offer limited possibilities to reason about representations, as visualisation tools were lacking on the visited websites, and only some sites offered star ratings distribution bar graphs.

5 Discussion and conclusion

5.1 Answer to the research question

In this observational multiple case study, we explored: *How do adults reason statistically with web-search results to answer daily life questions?* Adults' questions included finding products and mortgages. This under-researched area (Gal et al., 2020) has few empirical studies (Gal, 2024a). A key strength of our study is observing adults during their Web searches, alongside questioning them on their actions and reasoning.

The first part of the answer to the research question is that, although participants followed steps similar to those in the investigative cycle (Wild & Pfannkuch, 1999), their use of Web searches for consumer tasks (cf. Gal, 2000) impacted what data they found and how they reasoned statistically with them. Whereas reasoning about data (Biehler et al., 2018) spans all steps of the investigative cycle, we focused on the data-ing part related to the initial steps of this cycle including obtaining and preparing data for analysis. Critical evaluation and meaning-making (of the data) are important to these consumer tasks (Gal, 2000).

For the second part of the answer to the research question we looked at the data that participants encountered, their data-ing, and their statistical reasoning. Firstly, the data they found were unstructured (links presented as images with text, text links, calculators) and structured (prices, reviews in webshops) (cf. Cafarella et al., 2011). These data may vary in time and depend on randomisation and cookie settings (Urman & Makhortykh, 2023). Moreover, in web-search data, a clear data producer (researcher, institute) was often absent, the population was unclear, there was no single dataset, samples were non-random, how and why data were produced was often unknown, and data stemmed from multiple sources and were updated irregularly.

Secondly, we found several examples of data-ing. Participants evaluated data, applied data moves (Erickson et al., 2019), and used proxy variables (e.g., year of publication for relevance). Websites that list items (webshops, homes sites) provide many opportunities for ordering and filtering data but participants' skills in effectively using these options varied.

Thirdly, participants rarely reasoned explicitly with statistical concepts, but their reasoning reflected ideas such as association and variation. For example, they informally associated two variables (e.g., price and brand) in ways not fitting current descriptions of reasoning about association (e.g., Garfield, 2003). Aligned with literature on IIR (Zieffler et al., 2008), we use the term *informal statistical reasoning* to describe how adults use informal statistical knowledge to reason about statistical concepts such as measures, samples, associations, populations, variability, representations, bias, and uncertainty. Here, *informal* refers to adults' understanding based on daily life experiences and less structured knowledge acquired through earlier education, which lacks the formal rigour of academic statistical reasoning (cf. Zieffler et al., 2008).

5.2 Limitations

A limitation of our study is the small convenience sample, which may affect the generalisability to a broader sample of adults. Future research could explore whether similar data-ing and informal statistical reasoning are observed among adults with only primary or junior vocational education, as all participants in this study had senior vocational qualifications. Additionally, future studies could examine different search devices (e.g., mobile phones, tablets) and tasks to reveal other forms of statistical reasoning.

Despite attempts to mimic a realistic setting, participants performed tasks at our request, which may have affected their search behaviour. Future studies could include in-depth interviews to reveal misinterpretations now only touched upon. Future research should also explore how other sources (e.g., social media) and artificial intelligence affect adults' data-ing and informal statistical reasoning with contemporary data.

5.3 Contributions

5.3.1 Scientific contribution: data producers versus data consumers

Discussing our findings, we note several similarities and differences between participants' steps and those described by the investigative cycle (Wild & Pfannkuch, 1999). Some steps appear similar: participants started with a question (*problem*), searched for data (*data*), *explored* and *analysed* them, and drew *conclusions*. Other steps were clearly different, mainly when it comes to participants' data-ing. As outlined in the theoretical background, consumers' tasks differ from producers' tasks. Participants did not produce data (or visualisations, or statistical measures). Hence, they did not make a plan for data management and they did not decide how to operationalise variables (*plan*). Instead, data were provided by others like Google, online retailers, and social media. Therefore, they needed to choose (cf. Strohmayer & Muller, 2023) what part of the—often unstructured—data they considered relevant and trustworthy for their investigation. Furthermore, their data moves (Erickson et al., 2019) were driven by possibilities on websites instead of what results they could get from applying them to a downloaded dataset. In a similar vein, participants did not create data visualisations, quantify uncertainty, or calculate statistical measures (*analysis*) (cf. Garfield, 2003) but, instead, needed to interpret them (Gal, 2000). The findings suggest that the investigative cycle may need to be reconsidered or updated to better address adults' contemporary data inquiries, which is left for future research.

Our findings are in line with Gal's (2002) distinction between the activities of data producers and data consumers. While data producers engage in generating and structuring data (e.g., deciding on sampling methods or operationalising variables), our participants, as data consumers, were tasked with interpreting pre-existing, often unstructured, data. This resonates with Gal's (2000) emphasis on the different cognitive demands placed on consumers, who must critically assess the relevance and trustworthiness of data rather than produce them. Consequently, participants had to navigate the limitations of available data.

This study theoretically contributes to the discussion on what data-ing entails. For data producers, data-ing involves considering how and what data should be *produced*. In contrast, data-ing for contemporary data focuses on *choosing* relevant data (cf. Strohmayer & Muller, 2023) and reflecting on their quality and potential bias. In both cases, data-ing includes finding suitable proxy variables when necessary. As statistical reasoning is central to statistical literacy, we also aim to stimulate discussions on updating statistical literacy to reflect the contemporary data citizens now encounter.

5.3.2 Educational contribution: Rethinking the curriculum from a consumer perspective

Our findings lend support to existing calls made by Engel and Ridgway (2022) and Gal (2024b) to reconsider secondary and vocational statistics education curricula. Current statistics curricula predominantly focus on data-production tasks (e.g., Bargagliotti et al., 2020; Biehler et al., 2018; Garfield, 2003; Strohmayer & Muller, 2023). Although some efforts have been made to integrate data science practices (Fry & Makar, 2021; IDSSP, 2019) most students are more likely to engage with data as consumers rather than producers or data scientists. To ensure students are prepared to handle contemporary data, curricula should include activities that develop competence in critically filtering information and assessing the credibility of sources like sponsored links, fostering more educated data consumers.

We advocate enhancing students' statistical Web search skills using authentic contexts based on their interests (e.g., movies, games, make-up, the environment, deferred payment). We suggest placing emphasis on what conclusions can and cannot be drawn from Web-search results. Possible skills to teach include data selection, distinguishing between ads and other results, understanding search and randomisation in webshop ordering, recognising credibility indicators, and developing data filtering skills. To support this, we hope that our perspective on data-ing inspires teachers to address concepts such as proxy variables and the trustworthiness, reliability, bias, and relevance of contemporary data. Shifting the focus of curricula towards developing informed data consumers can help improve statistical literacy and prepare future generations to engage more critically with contemporary data, ultimately fostering a more data-literate society.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11858-025-01665-4.

Acknowledgements Deepl.com was used for translation (e.g., transcripts). We thank Tina Dulam, Vuslat Şeker, and Anita Alexander for their feedback, and Tina for conducting an interview.

Funding Funding was supported by the Ministerie van Onderwijs, Cultuur en Wetenschap.

Declarations

Conflict of interest The authors declare no conflict of interest.

Ethical statements HU ECO-SD ethical committee advised positively (reference 2022-3).

References

Arasu, A., & Garcia-Molina, H. (2003). Extracting structured data from web pages. In *Proceedings of the 2003 ACM SIGMOD international conference on Management of data* (pp. 337–348). https://doi.org/10.1145/872757.872799

Bakker, A. (2014). Characterising and developing vocational mathematical knowledge. *Educational Studies in Mathematics, 86*, 151–156. https://doi.org/10.1007/s10649-014-9560-4

Bargagliotti, A., Franklin, C., Arnold, P., Gould, R., Johnson, S., Perez, L., & Spangler, D. (2020). *Pre-K-12 guidelines for assessment and instruction in statistics education (GAISE) report II*. American Statistical Association and National Council of Teachers of Mathematics. https://www.amstat.org/asa/files/pdfs/GAISE/GAISEIIPreK-12_Full.pdf. Accessed 19 Feb 2025.

Ben-Zvi, D., & Garfield, J. (2004). Statistical literacy, reasoning, and thinking: Goals, definitions, and challenges. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 3–15). Springer. https://doi.org/10.1007/1-4020-2278-6_1

Biehler, R., Frischemeier, D., Reading, C., & Shaughnessy, J. M. (2018). Reasoning about data. In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education*. Springer. https://doi.org/10.1007/978-3-319-66195-7

Boels, L., Hoogland, K., Jonker, V., Kleine Deters, B., & Wijers, M. (2022) *Het interpreteren en begrijpen van hedendaagse informatiebronnen* [Interpreting and making sense of contemporary information sources]. (pp. 1–84). HU University of Applied Sciences Utrecht.

Boisot, M., & Canals, A. (2004). Data, information and knowledge: Have we got it right? *Journal of Evolutionary Economics, 14*, 43–67. https://doi.org/10.1007/s00191-003-0181-9

Cafarella, M. J., Halevy, A., & Madhavan, J. (2011). Structured data on the Web. *Communications of the ACM, 54*(2), 72–79. https://doi.org/10.1145/1897816.1897839

Corbin, J., & Strauss, A. (1990). Grounded theory research: Procedures, canons and evaluative criteria. *Zeitschrift Für Soziologie, 19*(6), 418–427. https://doi.org/10.1515/zfsoz-1990-0602

delMas, R. C. (2004). A comparison of mathematical and statistical reasoning. In D. Ben- Zvi, & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 79–95). Springer. https://doi.org/10.1007/1-4020-2278-6_4

delMas, R., Ooms, A., Garfield, J. & Chance, B. (2006). Assessing students' statistical reasoning. In A. Rossman & B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics*. IASE, ISI. https://iase-web.org/documents/papers/icots7/6D3_DELM.pdf?1402524965

Engel, J., & Ridgway, J. (2022). Back to the future: Rethinking the purpose and nature of statistics education. In J. Ridgway (Ed.), *Statistics for empowerment and social engagement.* Springer. https://doi.org/10.1007/978-3-031-20748-8_2

Engel, J., Sedlmeier, P, & Wörn, C. (2008). Modelling scatterplot data and the signal-noise metaphor: Towards statistical literacy for pre-service teachers. In C. Batanero, G. Burrill, C. Reading & A. Rossman (Eds.), *Joint ICMI/IASE Study: Teaching statistics in school mathematics. Challenges for teaching and teacher education. Proceedings of the ICMI study 18 and 2008 IASE Round table conference.* https://iase-web.org/documents/papers/rt2008/T4P8_Engel.pdf

Erickson, T., Wilkerson, M., Finzer, W., & Reichsman, F. (2019). Data moves. *Technology Innovations in Statistics Education*. https://doi.org/10.5070/T5121038001

Fry, K., & Makar, K. (2021). How could we teach data science in primary school? *Teaching Statistics, 43*, S173–S181. https://doi.org/10.1111/test.12259

Gal, I. (2000). Statistical literacy: Conceptual and instructional issues. In D. Coben, J. O'Donoghue, & G. FitzSimons (Eds.), *Perspectives on adults learning mathematics: Theory and practice* (pp. 135–150). Springer. https://doi.org/10.1007/0-306-47221-X_8

Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review, 70*(1), 1–25. https://doi.org/10.1111/j.1751-5823.2002.tb00336.x

Gal, I. (2024a). Adult education in mathematics and numeracy: A scoping review of recent research. *ZDM Mathematics Education*. https://doi.org/10.1007/s11858-024-01549-z

Gal, I. (2024b). What does it mean to prepare our graduates for adult life? Perspectives from statistics education and adult numeracy. *Presentation at the 15th International Congress on Mathematical Education, Sydney, Australia*.

Gal, I., & Geiger, V. (2022). Welcome to the era of vague news: A study of the demands of statistical and mathematical products in the COVID-19 pandemic media. *Educational Studies in Mathematics, 111*, 5–28. https://doi.org/10.1007/s10649-022-10151-7

Gal, I., Grotlüschen, A., Tout, D., & Kaiser, G. (2020). Numeracy, adult education, and vulnerable adults: A critical view of a neglected field. *ZDM Mathematics Education, 52*(3), 377–394. https://doi.org/10.1007/s11858-020-01155-9

Garfield, J. B. (2003). Assessing statistical reasoning. *Statistics Education Research Journal, 2*(1), 22–38. https://doi.org/10.52041/serj.v2i1.557

Garfield, J. B., Ben-Zvi, D., Chance, B., Medina, E., Roseth, C., & Zieffler, A. (2008). *Developing students' statistical reasoning: Connecting research and teaching practice*. Springer. https://doi.org/10.1007/978-1-4020-8383-9

Geiger, V., Goos, M., & Forgasz, H. (2015). A rich interpretation of numeracy for the 21st century: A survey of the state of the field. *ZDM Mathematics Education, 47*(4), 531–548. https://doi.org/10.1007/s11858-015-0708-1

Gigerenzer, G. (2022). Foreword. Democracy needs statistical literacy. In J. Ridgway (Ed.), *Statistics for empowerment and social engagement. Teaching civic statistics to develop informed citizens* (pp. v–vii). Springer. https://doi.org/10.1007/978-3-031-20748-8

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

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

Hoogland, K., & Díez-Palomar, J. (2022). The Mathematisation of society: Rethinking basic skills for adults. In J. Hodgen et al. (Eds.), *Proceedings of the twelfth congress of the European Society for Research in Mathematics Education, Bozen-Bolzano, Italy*. https://hal.archives-ouvertes.fr/hal-03745525v1. Accessed 19 Feb 2025.

Hoogland, K., & Stelwagen, R. (2021). *Het belang van rekenen en gecijferdheid in een levenlang ontwikkelen [The importance of numeracy in lifelong development]*. HU University of Applied Sciences Utrecht & Probiblio.

IDSSP. (2019). *Curriculum frameworks for introductory data science.* http://www.idssp.org/files/IDSSP_Frameworks_1.0.pdf. Accessed 19 Feb 2025.

Kanes, C. (1996). Investigating the use of language and mathematics in the workplace settings. In P. Clarkson (Ed.), *Technology in mathematics education* (pp. 314–321). Melbourne: MERGA. https://merga.net.au/wp-content/uploads/DOCS/common/Uploaded%20files/Annual%20Conference%20Proceedings/1994%20Annual%20Conference%20Proceedings/RP_Kanes_1996.pdf. Accessed 19 Feb 2025.

Kim, D., Park, Y., Yoon, M., & Jo, I. (2016). Toward evidence-based learning analytics: Using proxy variables to improve asynchronous online discussion environments. *The Internet and Higher Education, 30*, 30–43. https://doi.org/10.1016/j.iheduc.2016.03.002

Lewandowski, D., & Schultheiß, S. (2023). Public awareness and attitudes towards search engine optimization. *Behaviour and Information Technology, 42*(8), 1025–1044. https://doi.org/10.1080/0144929X.2022.2056507

Lovett, M. (2013). A collaborative convergence on studying reasoning processes: A case study in statistics. In S. M. Carver & D. Klahr (Eds.), *Cognition and instruction* (pp. 347–384). Taylor & Francis.

Necula, S.-C., & Păvăloaia, V.-D. (2023). AI-driven recommendations: A systematic review of the state of the art in e-commerce. *Applied Sciences, 13*(9), 5531. https://doi.org/10.3390/app13095531

OECD (2013). *Country note. Netherlands. Survey of adult skills first results*. OECD Publishing. https://www.oecd.org/content/dam/oecd/en/about/programmes/edu/piaac/country-specific-material/cycle-1/Netherlands-Country-Note.pdf. Accessed 19 Feb 2025.

OECD. (2024). Do adults have the skills they need to thrive in a changing world? Survey of adult skills 2023. *OECD Skills Studies*. https://doi.org/10.1787/b263dc5d-en

Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G., & Granka, L. (2007). In Google we trust: Users' decisions on rank, position, and relevance. *Journal of Computer-Mediated Communication, 12*(3), 801–823.

Ridgway, J., Nicholson, J., & McCusker, S., et al. (2011). Developing statistical literacy in students and teachers. In C. Batanero (Ed.), *Teaching statistics in school mathematics: Challenges for teaching and teacher education* (pp. 311–322). Springer. https://doi.org/10.1007/978-94-007-1131-0_30

Rufiana, I. S., Sa'dijah, C., Dubanji, Susanto, H., & As'ari, A. R. (2018). Informal statistical reasoning of students taken formal statistics learning related to distribution. *International Journal of Insight for Mathematics Teaching, 1*(2), 130–140.

SRTL (2022). *SRTL13 preliminary announcement. The Thirteenth International Research Forum on Statistical Reasoning, Thinking and Literacy (SRTL-13)*. Retrieved February 19, 2025, from https://blogs.uni-paderborn.de/srtl/files/2022/09/SRTL13-Preliminary-announcement.pdf.

Strohmayer, A., & Muller, M. (2023). Data-ing and un-data-ing. *Interactions, 30*(3), 39–42. https://doi.org/10.1145/3587240

Strzelecki, A. (2020). Eye-tracking studies of web search engines: A systematic literature review. *Information, 11*(6), 300. https://doi.org/10.3390/info11060300

Tracy, S. J. (2013). *Qualitative research methods*. Wiley.

Twining, P., Heller, R. S., Nussbaum, M., & Tsai, C.-C. (2017). Some guidance on conducting and reporting qualitative studies. *Computers & Education*, *106*, A1–A9.

Urman, A., & Makhortykh, M. (2023). You are how (and where) you search? Comparative analysis of web search behavior using web tracking data. *Journal of Computational Social Science*. https://doi.org/10.1007/s42001-023-00208-9

Wild, C., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review, 67*(3), 223–248. https://doi.org/10.1111/j.1751-5823.1999.tb00442.x

Wise, A. F. (2020). Educating data scientists and data literate citizens for a new generation of data. *Journal of the Learning Sciences, 29*(1), 165–181. https://doi.org/10.1080/10508406.2019.1705678

Zieffler, A., Garfield, J., delMas, R., & Reading, C. (2008). A framework to support research on informal inferential reasoning. *Statistics Education Research Journal, 7*(2), 40–58. https://doi.org/10.52041/serj.v7i2.469