

**AN ANALYSIS OF EFFICIENCY OF DATA COMMUNICATING: EXPLORING THE
IMPACT OF DATA TYPES ON THE CONSUMERS' COMPREHENSION SCOPE**

YUNPING CAI

Thesis submitted to the University of Ottawa
in partial Fulfillment of the requirements for the
Master of Arts

Department of Communication

Faculty of Arts

University of Ottawa

© Yunping Cai, Ottawa, Canada, 2024

Abstract

This study ($N = 46$) employed a multimodal approach, combining surveys and eye-tracking technology, to explore factors influencing audiences' comprehension of data found in a news article. The study sought to investigate the connections among participants' dispositions, the presentation methods of data types, audience attention spans, and the comprehensiveness of audience understanding. Leveraging Gal's (2002) data literacy model, the research explored the integration of literacy and dispositional elements and their impact on data literacy within the context of print news. Findings indicated a significant positive correlation between motivation, career aspirations, and data comprehension levels, suggesting a future-oriented pattern in consuming data in news. Additional findings further illuminate the gender gap in numeracy with respect to dispositional factors: while men excel in mental calculation, women are more adept at graph reading and information retrieval. These dispositional factors also contribute to the observed gender differences: motivational factors significantly influence men's perception of difficulty, whereas career aspirations play a pivotal role in enhancing women's comprehension abilities, leading them to perform better than men. The study highlights the need to consider evidence-based best practices for designing quantitative information in news, considering how audience disposition may influence the complexity, types, and organization of this information to enhance communication effectiveness. Such strategic presentation can significantly empower audiences, fostering informed decision-making and enhancing the quality of rational public discourse.

Keywords: Data literacy, data in news, data communication, data journalism, quantitative reasoning, consumer reception, path analysis, scientific reporting, eye tracker,

Table of Contents

Abstract	ii
Chapter 1 Introduction	1
Chapter 2 Background	4
2.1 Data Categories: Presentations and Types	4
2.2 Data Literacy.....	10
2.3 Dispositional Challenges in Data Literacy.....	14
2.4 Limitations and Proposed Research.....	16
2.5 Research Questions	19
Chapter 3 Research Design and Methodology	21
3.1 Research Strategy.....	21
3.11 Variable Design	21
3.12 Disposition Assessment	22
3.13 Participant Engagement via Eye-Tracking Technology.....	23
3.14 Data Comprehension.....	24
3.2 Methodology.....	25
3.21 Procedures.....	25
3.22 Participant Recruitment.....	27
3.23 Participants.....	28
3.24 News Article Selection.....	30
3.25 Measures	31
3.26 Analytical Plan.....	42
Chapter 4 Results and Analysis	46
4.1 Data Cleaning and Selection.....	46
4.2 Descriptive Statistics.....	47
4.21 Demographics and Disposition	47
4.22 Eye-Tracking Results of News Article	49
4.23 Data Comprehension.....	53
4.3 Participant Background.....	56
Gender.....	57
English As First Language.....	65

Parental Education Level	67
4.4 Correlations.....	68
Relationship Between Attention Span and Data Comprehension Performance: Dependent Variables 1 and 2 (DVs #1 & DVs #2)	69
Relationship Between Dispositions and Data Comprehension Performance: Independent Variables (IVs) and Dependent Variable 2 (DVs #2)	70
Relationship Between Dispositions and Attention Spans: Independent Variables (IVs) and Dependent Variable 1 (DVs #1).....	71
4.5 Predictive Model.....	73
Regression.....	74
Path Analysis	83
Chapter 5 Conclusion	92
5.1 Findings.....	92
RQ1: Data Engagement	92
RQ2: Engagement, Comprehension, and Perceived Difficulty.....	93
RQ3: Participant Dispositions, Comprehension, and Perceived Difficulty	94
Additional Findings	95
5.2 Discussion.....	97
5.2 Improvement.....	99
5.21 Methodological Limitations.....	99
5.22 Theoretical Limitations.....	101
References.....	104
Appendices.....	114

Chapter 1 Introduction

In this data-dominated era, the escalating impact of data on shaping public perception underscores the importance of data as a crucial commodity, a contemporary communication method, a cross-cultural dialogue, a universal metric, and a rational instrument for democratic engagement. In recent years, the media has transformed into a quantitative forefront, reporting on data coming from financial institutions, think tanks, governmental bodies, and academic research. This quantitative shift supports informed decision-making and justifies governmental interventions, as evidenced during the COVID-19 pandemic (Gal & Geiger, 2022). This era of data abundance promotes rational discourse, marking a shift in media practice to a more data-centric approach (Borges-Rey, 2016).

Data, or quantitative information, in news typically involves numerically expressed details like statistics and data visualizations, that support narratives within journalistic reports. This information is mostly statistical, focusing on conveying and interpreting data, rather than on mathematical aspects like calculations or model building (Nguyen & Lugo-Ocando, 2016). However, effectively understanding this information demands that audiences distinguish credible sources and assess future risks—a skill that is complex and not as straightforward as it might seem. Successfully conveying quantitative information requires data literacy at both the media's end for accurate presentation and at the audience's end for effective comprehension (Koltay, 2015).

The critical, yet underappreciated role of quantitative information in news highlights a gap in media discussions about data literacy. Strategies for conveying quantitative information

frequently center around the practice of data journalism, which is viewed as a way to create news gathering techniques that can democratize the flow of information (Bradshaw, 2011; Rogers, 2011). Data journalism, which encompasses everything from creating infographics to analyzing raw data sets, plays a central role in how journalists adapt to the data-driven era (Knight, 2015). Current research practices focus on journalists' roles in (1) validating the source of information, whether it involves high-precision scientific statistics or broader socio-political data (e.g., (Kovach & Rosenstiel, 2014), (2) enhancing journalists' professional understanding of statistics and data (e.g., Cushion et al., 2017), and (3) effectively communicating complex data in ways that accommodate varying levels of audience data literacy (e.g., Cushion et al., 2017). Despite these advancements, the primary focus remains on the media's misuse or mishandling of data, neglecting the need for designs that cater to different consumer literacy levels, and what's more, what could foster or improve their literacy. This oversight points to the necessity for a more profound examination of the building blocks of data literacy, as Gal (2002) suggests, to help individuals navigate the complex quantitative demands of media consumption, as well as everyday life.

This study addresses the existing gap between media practices and audience comprehension by exploring how news consumers interact with and understand quantitative information they encounter while reading a news article. Utilizing a lab environment, it builds on Gal's (2002) data literacy model to develop a framework for data literacy in news, examining how demographic, dispositional, and literacy factors influence the reception and comprehension of news data. The exploratory research also incorporates various formats of quantitative information, using eye-tracking technology to measure engagement and establish links between

disposition, presentation format, and audience understanding. This approach calls for an integrated, interdisciplinary research methodology that combines psychological, social science, and communication studies to delve into effective data literacy in news.

Chapter 2 Background

An extensive body of research has been conducted to investigate the effectiveness of how quantitative information is communicated and the audience's reception of quantitative elements in news media. This existing research encompasses the categorization of news articles, coding and detailed classification of the quantitative elements within, and has ultimately delved into the audience's reception, regardless of their statistical literacy levels. The majority of these studies have primarily focused on generalizing how news pieces are constructed and how the data presented within the articles influences audience comprehension, while the interplay between them remains a gap.

The following literature review will start with an exploration of the prevailing frameworks that have shaped the comprehension of audience reception of quantitative information within news media. It will then refine the focus to delve into the meticulous categorization of quantitative information.

2.1 Data Categories: Presentations and Types

To begin the theoretical review, it is imperative to provide a precise and contextual definition of 'data' within the realm of printed news reporting in order to establish a solid academic framework. In journalism, 'data' typically refers to a collection of factual, statistical content that underpins news stories, providing a reliable foundation for reporting (Nguyen & Lugo-Ocando, 2016). Additionally, it is important to view data not just as raw numbers but as numbers contextualized within a framework that adds meaning and relevance (Gal, 2002). It manifests in

multiple formats: textual accounts, raw figures, mathematical constructs, and varied visualizations. These formats serve not only to impart facts, but also to prompt the audience's interpretive, analytical engagement and valid reasoning beyond mathematical calculation with the material (Gal & Geiger, 2022). In this context, 'data in news' signifies the mathematical, numerical and statistical content disseminated in printed news media.

News data can be organized in many different formats and considers two key components: 1) data types, focusing on the inherent characteristics of mathematical and numerical entities; and 2) data presentations, examining the cohesive relationships between data and its contextual roles within the news article.

Data types can be described according to Gal & Geiger's (2022) categories of Statistical and Mathematical Products (StaMPs) in media and Barchas-Lichtenstein et al.'s (2022) method for developing quantitative epistemologies. These include five distinct types: (1) Proportion or percentage, (2) Statistical statements (average, mean, central tendencies and exceptions), (3) Data relationship (correlation and comparison modelling), (4) Absolute numbers (raw numbers without context), and (5) Contextualizers (age, date, gender, socio-economic status, etc.).

Data presentation encompasses three distinct formats through which data is conveyed in printed news media: (1) Narrative style encompassed in sentences (Ojo & Heravi, 2018), (2) Static data visualization (graphics, maps and charts) (Maltese et al., 2015), and (3) Raw and isolated data. Isolated numbers are rare in news articles (Whitacre, Henning, & Atabaño, 2020), instead data is usually presented alongside contextualization or explanations. Often, figures provide a certain

conceptual framework surrounding the data itself, which guide interpretation of predictive models and inductive reasoning, such as probability, magnitude, central tendencies, and variability (Barchas-Lichtenstein et al., 2022).

The integration of data into news reporting is recognized as a strategy for enhancing journalistic rigor as numbers are perceived as trustworthy, objective and value-free (Rogers, 2011). Notably, no less than 44 percent of news articles include data such as percentages, amounts, or proportions (Zillman and Brosius, 2012). Some preliminary research has begun to show that visual presentations may be more effective than narrative or raw data presentations as they reduce cognitive processing and promote comprehension and recall (Mahmoud & Zoghaib, 2023; Sabramanya, 2016). However, although there is some evidence that certain data presentation formats are more effective than others, these findings do not apply to all types of data within all contexts. For example, one study found audience comprehension and recall improved by contextualizing data with short textual analogies (Barrio et al., 2016).

Data, numerical representations, scientific tools, and technology are intricately interconnected, with data itself perceived as a technological entity possessing versatile utility, and notably, deemed fundamentally neutral in moral stance (Winner, 1986). This inclination towards authenticity and objectivity is underpinned by the notion that a positive correlation exists between technological progress, objectivity, democracy, and ethical considerations (Winner, 1986). Because people do not generally engage with raw data, journalists often play a key role in tracking down, interpreting and presenting numerical data (Cushion et al., 2017). Journalists can promote audience understanding by strategically selecting how they present the data.

The emergence of data journalism reflects a growing demand for credibility and objectivity in how audiences consume news. Implicit in this pursuit of objectivity is a commitment to democratic ideals, with the ascent of modern data journalism being deeply intertwined with political motives, particularly the open data movement, which advocates for an information democracy (Chao, 2023). While the era of data transparency promises a future of open information sharing, the capacity of individuals to utilize this data for informed decision-making remains an area of uncertainty.

First, the neutrality and the positive implications associated with this neutrality, as proposed by quantitative information, shape non-verbal communication pieces in both public discourse and personal worldview building. According to Merriam (1990), numbers have persuasive effects because they invoke deeply embedded connotations and beliefs that numbers are better than words. Koetsenruijter (2011), in an empirical study examining the use of data in news articles for credibility, found a relationship between credibility assessments of news consumers and the presence of numbers. This finding has noteworthy practical implications for journalists as the strategic use of numbers in their articles may shape and positively influence audience perception.

The delicate relationship between the assumed neutrality of numbers, statistical narratives, and mathematical models (Jablonka & Bergsten, 2021) (referred to in this study as "data in news") and their ability to persuade the public towards compliance and consensus, whether concerning controversial policies or scientific education, underwent a critical assessment by Jablonka and Bergsten (2021). Using a content analysis under the post-structuralist discourse theoretical

framework, their study identified four strategies—rationalization, contrast, association, and recharging—that were linked with manipulating the selection and adaptation of data in news to serve particular interests. Although their results highlighted the complexity of how numbers are used in public discourse, the study lacked a model explaining the interplay between audience reception and data in news. Roeh and Feldman (1984) termed the characteristics and mechanisms employed by data to persuade audiences as the "rhetoric of numbers." They concluded that "the function of numbers in the press may be to create an impression of facts-and-only-the-facts". According to their perspective, numbers are utilized more frequently for their rhetorical impact than for their capacity to provide factual information. Lindsey and Yun (2003) conducted a study examining the relationship between the use of statistical evidence in a persuasive appeal and an individual's attitude. Their findings align with the notion that "data in news affects attitude" where individuals respond favourably to articles containing data. They also highlighted the significance of the sample size heuristic (i.e., individuals' inability to consider sample size when interpreting and generalizing findings) as a major determinant of attitudes when individuals are exposed to statistical messages. The conclusions drawn by these critics regarding the neutrality of data in news underscore the necessity for a thorough and revalidated scrutiny of the role played by data in news media content.

Additionally, the tolerance for ambiguity or even contradictions within data in news is influenced by factors such as timeliness, credibility of sources, numerical competence of reporters, and audiences' critical mathematical literacy, reading literacy, and quantitative reasoning (referred to in this study as "data literacy"). In Robert Brand's (2008) case study assessing the numerical accuracy of a South African newspaper, he posited that a lack of proficiency or attention to basic

arithmetic significantly contributed to errors in quantitative elements within the news, thereby compromising democratic discourse and influencing citizens' ability to make informed decisions. According to Brand, ambiguity poses a threat to the news media's role as the "watchdog," as errors in data in news present a potential avenue for misleading and manipulating public discourse in favor of specific interests (Brand, 2008). While Brand's apprehensions were addressed through his content analysis of quantitative elements in news, a gap remains regarding the potential effects of these data errors in increasing ambiguity and contradictions, and the audience's capacity to identify the errors. This gap may necessitate future examination to validate the relationship between Brand's arguments and the potential impact of data errors on public understanding and decision-making processes.

Another investigation conducted by Theivendrampillai et al. (2023) examining misinformation in COVID-19 reporting by news media revealed that misinformation, information overload, or the absence of information could lead to conflicting public acceptance of policies related to vaccines and contribute to stigmatization around racialized issues. Following challenges that were already emerging before the COVID-19 pandemic, Cushion et al. (2017) proposed a solution in their study focused on statistical information in United Kingdom news pieces. They suggested that independent scrutiny of statistics at the newsgathering and journalistic levels serves as a means to address contradictions. According to Cushion et al. (2017), critical judgments should be made with objectivity in statistical reporting, navigating the complex terrain of impartiality and drawing information from a broader range of sources. A comprehensive comparison and review of existing studies on perspectives and data neutrality, along with the implications for democracies stemming from the intrinsic objectivity of quantitative information, has uncovered a

gap in research. This gap necessitates an inquiry into whether data in news media authentically embodies neutrality and enhances the audience's comprehension of news media information during consumption.

In conclusion, data presented in the news is not inherently neutral or objective. Its presentation and interpretation are shaped by various factors, including journalists' and audience-members' dispositional factors. Data literacy, which encompasses a broad range of literacies including media and statistical literacy, is critical in determining how individuals contextualize, interpret, and understand data in their daily lives. According to Brand (2008), the data literacy of reporters and journalists affects their ability to understand statistical models and to incorporate quantitative information effectively into news pieces, thereby enhancing persuasiveness and clarity. Similarly, the data literacy of audiences significantly influences how they interpret and comprehend the data presented in news reports.

2.2 Data Literacy

Gal (2002) introduced a comprehensive framework (originally defined as Statistical Literacy, but later redefined as Data Literacy) for describing data literacy in the context of news media and everyday citizenship. Gal's model delineates five cognitive or knowledge elements—literacy, statistical, mathematical, context, and critical—and two dispositional ones—beliefs/attitude and critical stance. According to this framework, these are necessary for the audience's ability to a) interpret and critically evaluate statistical information and b) articulate their reactions, encompassing understanding, opinions, and concerns within diverse contexts, constituting adult statistical literacy (Gal, 2002). Gal emphasizes the significance of statistical literacy, not only in

heightening citizens' awareness in the information era but also in enriching democratic discourse within the public sphere.

Figure 1: A model of datal literacy (Gal, 2002)

Knowledge Elements	Dispositional Elements
Literacy skills Statistical knowledge Mathematical Knowledge Context Knowledge Critical Questions	Beliefs and Attitudes Critical stance
Data Literacy	

Note: This figure represents the components of statistical literacy as conceptualized by Gal (2002), encompassing both knowledge and dispositional elements.

Data literacy, as defined by Gal, involves several key components:

1. **Literacy Skills:** The ability to process text to derive meaning from data presented to readers.
2. **Statistical Knowledge:** Understanding basic statistical concepts, including the necessity and production of data, familiarity with descriptive statistics, and interpreting graphical and tabular displays.
3. **Mathematical Knowledge:** Recognizing that statistical summaries (e.g., means, percentages) represent a concise way to describe large data sets and an understanding of the mathematics behind statistical indicators.
4. **Contextual Knowledge:** Insights into the background of the study or data, which aids in making sense of the information presented.

5. **Critical Questions:** This component involves the ability to scrutinize the reasonableness of claims and the validity of the evidence presented in the media. It emphasizes engaging in a complex and challenging thinking process, motivated by a desire to understand and evaluate alternative interpretations.
6. **Critical Stance:** This reflects a proactive, questioning attitude toward quantitative messages that may be misleading or biased. It involves motivated action, demonstrating a readiness to address and resolve uncertainties encountered in data presentations.
7. **Beliefs and Attitudes:** Personal feelings and culturally influenced beliefs that develop over time regarding various objects, actions, or topics.

Gal's model builds upon two existing frameworks that have examined how audiences' dispositional factors influence their consumption of news items, particularly in terms of information dissemination and the formation of attitudes towards shared content (Koetsenruijter, 2011). First, the Data-Driven Decision-Making (DDDM) model proposed by Ikemoto and Marsh (2007) and the concept of Quantitative Reasoning (QR) in news consumption as outlined by Karaali et al. (2016). These conceptual frameworks are crucial for analyzing how audiences engage with and interpret quantitative information within the media landscape.

DDDM emphasizes replacing subjective judgment with evidence-based practices that are grounded in data, transforming raw data into actionable knowledge. However, there is a noted lack of application of DDDM in real-life contexts. QR is defined as the capability to interpret numbers within context and to make informed decisions based on numerical information (Karaali et al., 2016). In Attaway et al.'s (2023) study on QR skills and news habits, QR skills are

conceptualized as "reasoning about and engaging with quantitative information," which extends beyond mere numerical operations or working with specific data representations like graphs (Attaway et al., 2023) in news items.

In contrast to Ikemoto and Marsh's DDDM model, which centers on actively learning statistics, Gal's data literacy model directs attention to adults and their roles as consumers in social lives, including everyday communication, community engagement, and news consumption. Unlike DDDM, data literacy places less emphasis on converting raw data into actionable knowledge but rather integrates multivariate elements into its conceptualization. Compared to Quantitative Reasoning (QR), data literacy is more theoretical and non-utilitarian, providing a theoretical framework across interdisciplinary landscapes. Furthermore, while data literacy (Gal, 2002) is perceived under various contexts as a fusion of basic statistical concepts and procedures, QR skills explicitly incorporate other prerequisite areas of mathematics (Barchas-Lichtenstein et al., 2021). This makes QR more focused on mathematical and numeracy literacy rather than factual knowledge and specific formal or informal skills, encompassing desired beliefs, habits of mind, attitudes, general awareness, and a critical perspective (Gal, 2002).

In essence, QR places emphasis on statistical concepts for decision-making without necessarily engaging in numerical operations or working with specific data representations such as graphs (Attaway et al., 2023), DDDM focuses on the transition from raw numbers to decision-making by converting raw data into actionable knowledge and information (Ikemoto and Marsh, 2007). In this context, data literacy integrates these two concepts into a unified model. This approach, which considers interconnected knowledge bases and dispositions (Gal, 2002), facilitates the

cognitive transformation from raw quantitative information to decision-making. It strategically addresses the limitations observed in previous studies conducted under the domains of Data-Driven Decision Making (DDDM) and Quantitative Reasoning (QR) research.

In summary, Data Literacy, DDDM, and QR all compel individuals to critically engage with numerical data and employ mathematical reasoning to derive meaningful conclusions from quantitative inputs. Fundamentally, all models require individuals to be motivated to engage with quantitative data in order to utilize it effectively for decision-making and critical reasoning. However, there are many dispositional factors that prevent audiences from doing this and one important challenge area is the relatively low numeracy levels among North American adults.

2.3 Dispositional Challenges in Data Literacy

Effectively reading, comprehending, reasoning, critically reflecting, and making decisions with data in news require complex competencies and skills. However, there is limited understanding of the mechanisms behind the interplay of dispositional factors in this process. For many news audience members, there are deficiencies and structures that prevent them from developing the skills necessary for interpreting and interacting with the data they encounter in the news.

One deficiency is deeply embedded within the North American mathematical education system, where prevalent social cues perpetuate the stereotype that if you struggle with math, you are inherently bad at it. Such perceptions significantly influence adults' willingness to develop their mathematical and statistical skills later in life, especially if they experienced discouragement during their formal education years (Beilock & Maloney, 2015). The pervasive stereotypes that

categorize individuals as "math" or "not-math" people significantly impact their ability to comprehend and make decisions based on data. This correlation was substantiated in Barroso et al.'s (2021) study, where a meta-analysis examined the updated relationship between math anxiety and math achievement, revealing that high math anxiety often coincides with low math achievement (Barroso et al., 2021). Moreover, the existing imbalance in numeracy education is further exacerbated by inequalities related to gender, racial disparities, and differences between communities (Barchas-Lichtenstein et al., 2021).

The current study acknowledges gender and age disparities in numeracy and news media literacy skills, with some discrepancies attributed to gender bias and stereotypes (Thomas, Hopwood, & Hatisaru, 2024; Cook, 2018; Krieg, Parada-Cabaleiro, & Schedl, 2022). Contrarily, Oyeniyi (2013) found no differences in media literacy and information retrieval skills across genders. Regarding numeracy, the gender differences cannot solely be explained by educational disparities. Instead, broader societal factors play a significant role in shaping these inequalities. These include unequal access to power and resources and gender stratification, as mentioned by Raude et al. (2023) in their exploration of the factors behind adult literacy differences. Another study examining the gender gap in adult numeracy emphasized that personality traits have a greater impact on women's numeracy skills than on men's. Lunardon et al. (2021) explored the complex interplay between math anxiety, neuroticism, and gender, finding that neuroticism significantly affects women's numeracy skills, whereas it does not have a comparable impact on men's. The emphasis on this gap highlights the challenges in organizational or strategic contexts and the ability to comprehend, analyze, and interpret numerical information across various scenarios due to dispositional differences.

Additionally, the use of data in careers varies; traditional views of labor divisions often assume that some jobs require numeracy while others do not. This division, influenced by industrialization, categorizes roles into "math-needed" and "math-not-needed" sectors, affecting adults' Data-Driven Decision Making (DDDM) and Quantitative Reasoning (QR). Factors like varying exposure to quantitative information, working memory (Ramirez et al., 2016), and motivation in mathematics (Wang et al., 2015) further shape these skills.

In terms of news habits and media literacy, studies such as those by Attaway et al. (2013) show a negligible link between news consumption patterns and quantitative reasoning. These findings suggest that while mathematical dispositions like anxiety and skill level are examined in relation to news media literacy, their impact on data literacy may be limited.

To summarize, there are many complex factors influencing consumers' understanding of data in news, data literacy emerges as an overarching concept, integrating DDDM, QR, and dispositional factors. What is missing from the current research repository is a comprehensive examination that spans from one end—mathematical and numeracy disposition—to the other: news items containing quantitative information or data literacy levels in news. Addressing this multifaceted interplay represents a significant gap in current studies.

2.4 Limitations and Proposed Research

The review of current theoretical models reveals a significant limitation: we lack a clear understanding of how audiences interact with data in news and comprehend the information

presented. Furthermore, while previous research has highlighted the role of dispositional factors such as numeracy, social cues, and stereotypes on mathematical competency, there has been insufficient detailed analysis to examine how these elements contribute to the effective reception, interpretation, and understanding of data in news. Additionally, no studies have yet controlled for audiences' reading patterns to confirm how they engage with data in news articles and to assess how these engagements serve as mediators in the effectiveness of quantitative information transmission to consumers.

As such, a data literacy framework that integrates key concepts and findings from the fields of news media, data journalism, and audience engagement will be used to guide the present study. This framework suggests that audience comprehension of news content is significantly influenced by the way data is presented within news stories and the cognitive abilities and dispositions of the audience.

Figure 2: Data Literacy Model

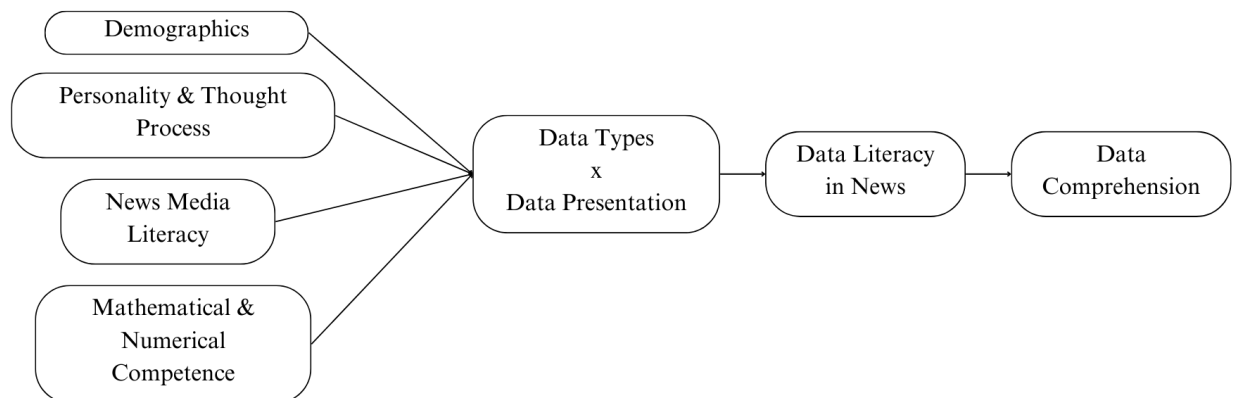
Dispositions	Data Types/Presentations		
<ul style="list-style-type: none"> • Demographics • Personality • Motivation and Critical Thinking • News Media Literacy • Mathematical and Numerical Competence 	Types: <ul style="list-style-type: none"> • Absolute Numbers • Contextualizers • Proportions and Percentages • Relationships and Comparisons • Statistical Statements 		
	Presentations:		
	Narrative	Raw	Visualization
Data Literacy in News			
Comprehension Level			

The framework encompasses several core components:

- **Data Presentation in News Media:** Different forms of data presentation, such as narrative data, visualizations, and raw data, are hypothesized to have varying effects on audience engagement and comprehension. The literature suggests that the complexity and accessibility of data presentations can either facilitate or hinder audience understanding.
- **Audience Cognitive Abilities and Dispositions:** The audience's ability to process and interpret data is influenced by their cognitive abilities, such as numeracy and literacy, as well as psychological factors like motivation, critical thinking, and openness to information. These individual differences are critical in determining how effectively audiences can engage with and comprehend data-driven news content.
- **Impact on Audience Engagement and Comprehension:** At the heart of the framework is the interaction between data presentation and audience dispositions, and how these factors collectively influence audience engagement with and comprehension of news content.

This theoretical framework serves as the foundation for the study's exploration into the dynamics of news media in the age of data journalism, aiming to uncover the factors that drive audience engagement and comprehension and to identify strategies for improving the effectiveness of news reporting in conveying complex information.

Figure 3: Proposed Study Model



Note: The framework bifurcates into 'Dispositions'—encompassing demographics, personality traits, motivation and critical thinking, news media literacy, along with mathematical and numerical competence—and 'Data Types/Presentations', categorized into 15 distinct sub-categories. These elements collectively influence the 'Comprehension Level' of data literacy within news media.

2.5 Research Questions

Building on the identified limitations regarding how audiences interact with data in the news and the role of dispositional factors in this process, this study seeks to respond to the following research questions:

RQ1: How do different data presentations or types impact audience attention and engagement with data presented in news articles?

RQ2.1: Does a relationship exist between the level of audience comprehension and the different data presentations or types featured in print news stories?

RQ2.2: Does a relationship exist between the level of audience perceived difficulty and the different data presentations or types featured in print news stories?

RQ3: Is there a connection between the levels of audience comprehension and their dispositional factors?

Chapter 3 Research Design and Methodology

3.1 Research Strategy

To address the research questions, this research employed an exploratory design within the context of a controlled laboratory study. Specifically, this study aimed to examine how participants' background dispositions as well as the presentation types of data found within a news article were related to participants' engagement and comprehension of the data within the article. Participants' dispositional backgrounds and comprehension were collected via self-report survey, whereas participants were equipped with an eye-tracker (Tobii Pro Spark) while reading a news article to assess engagement. Participant engagement was quantified by whether participants fixate on the data in the news article and for what duration (Tobii Academy, 2024), while comprehension is assessed based on their ability to correctly interpret the data. This study design allowed for a precise evaluation of participant interaction with each instance of data within the article.

3.11 Variable Design

The present research organized variables in accordance with the theoretical framework, categorizing participants' self-reported dispositions obtained from survey data as independent variables (IVs). Whereas engagement while reading news data, quantified by the time spent fixating on the different data as measured by Tobii eye-tracker, and data comprehension outcomes determined through post-reading response accuracy, were established as dependent variables (DVs). The rationale behind the selection of each of the measures are justified below.

3.12 Disposition Assessment

As the study's theoretical model indicates, participants' dispositions, including demographics, personality, thought processes, news media literacy, and mathematical and numerical competency, serve as independent variables contributing to data comprehension performance. To achieve this, an integrated survey was developed based on the research questions' requirements, assessing, measuring, and coding participants' dispositions. Survey items were selected, modified, and developed, drawing on existing survey items (John & Srivastava, 1999; Guay, Blais, Vallerand, & Pelletier, 1999; Ashley et al., 2013; Fagerlin et al., 2007; Nadratowski et al., 2022) in order to create valid and reliable variables representing the desired constructs.

Organized into four distinct measurement blocks, the study aimed to capture the following dispositional elements:

1. **Demographic Features:** This block captured essential demographic information, including gender, proficiency in English as a first language, and parental educational levels.
2. **Personality and Thought Processing:** Participants' personalities and general motivation were evaluated using the Big Five Inventory (BFI) and the Self-Determination Index (SDI). Additionally, thought processing (TP) was assessed to understand cognitive styles and complexities in problem-solving.
3. **News Media Literacy:** This segment evaluated participants' news media literacy, which includes News Media Knowledge (NMK) and News Media Literacy (NML). It included additional aspects such as Media Attitude, Media Experience, Media Confidence, and

Media Consumption, providing a holistic view of participants' engagement and perceptions of the media.

4. **Mathematical and Numerical Competence:** This block measures participants' competence in mathematics and numeracy, including the Subjective Numeracy (SN) and objective Numeracy (N) assessments. It also explores engagement with mathematics through indicators like math courses taken previously, average math grade achieved, and career aspirations in statistical and mathematical fields.

3.13 Participant Engagement via Eye-Tracking Technology

Eye-tracking technology offers a promising avenue for investigating incidental news consumption, providing objective data, free from biases like social desirability during the interpretation (Bucher & Schumacher, 2006). It enables the measurement of selective mechanisms and offers a means to bridge the gaps often found in self-reported outcomes due to participant fatigue or other uncontrollable factors (Bucher & Schumacher, 2006; Sülflow et al., 2019). Eye-tracking not only yields metrics of overt attention but also subscribes to the eye-mind assumption proposed by Just and Carpenter (1980), which posits that gaze fixations—resting the gaze on a stimulus—denote an individual's attention to the information presented. Such fixations are thought to indicate elaborate information processing (Loftus, 1972; Wedel & Pieters, 2000). However, the positive relationship between fixation duration and specific cognitive processes remains under-explored, leaving a gap for researchers to interpret the interactions between eye-tracking data and cognitive outcomes (Vergara et al., 2021).

It is crucial to recognize that eye movements are not merely the result of an automatic sensory mechanism, but are intricately linked with a person's actions, serving as tools for exploring the environment guided by individual intentions (Joos, Rötting, & Velichkovsky, 2003).

Consequently, this study advances existing research by examining how audiences engage with data in news stories through reading patterns captured by eye-tracking. It seeks to understand comprehensive post-reading behaviors, including quantified cognitive performances, a domain where little research has been conducted to integrate observational data with cognitive outcomes. This study acknowledges that patterns of interaction with media interfaces are shaped by both the interface design and the reader's expectations (Bucher & Schumacher, 2006). Therefore, eye-tracking data require careful examination and will be considered alongside other metrics to draw conclusive findings.

3.14 Data Comprehension

Data comprehension (DC) levels were assessed through open-ended questions derived from the data content of the article participants read. These questions required participants to understand the data presented in the article. To further enrich the dataset, each DC question was paired with a query about its perceived difficulty (DC Difficulty), thereby gauging the subjective challenge each question posed to participants.

The design and sequencing of these DC questions were planned to fulfill two objectives: 1) to integrate the influence of independent variables within the measurement process, and 2) to maintain a balanced level of difficulty across questions to prevent participant fatigue due to overly challenging content. After careful consideration, including a rearrangement of the

question sequence and the adoption of a coding scheme that allows for minor errors, this approach to measuring data comprehension was crafted. This strategy was designed to capture the levels of understanding that participants exhibit across various types of data presentation.

3.2 Methodology

3.2.1 Procedures

The exploratory study was conducted in the University of Ottawa Triple-I (Inclusion, Interaction, and Intervention) Lab, led by the Telfer-Arts Research Team. Upon arrival at the Triple-I Lab study venue, participants were guided to stations outfitted with desktop computers and eye-tracking devices and completed three distinct activities.

1. Dispositional Questionnaire: Participants completed a questionnaire via tablets using the Qualtrics Survey platform (Qualtrics, 2024). This questionnaire, taking approximately 15 minutes, assessed participants' dispositions, covering demographics, educational background, personality traits, cognitive processes, mathematical literacy, and attitudes towards media and news.

2. News Article Eye-Tracking Session: The Tobii Pro Spark eye-tracker, equipped with features such as a sampling frequency of up to 60 Hz, a single eye-tracking camera, and dual pupil tracking modes for enhanced accuracy in both bright and dark environments (Tobii Pro, 2024). The eye-tracker, integrated with Tobii Pro Lab software for sophisticated data analysis, was placed beneath each desktop monitor and the computer included peripherals like keyboards

and mice to ensure that participants could independently manage their progress and peruse the article at a self-directed pace. First, participants were calibrated to the eye-tracker. Calibration ensures precise eye movement tracking during the article reading and involves participants following dots on a screen to measure the accuracy of their gaze angles, with the mean offset in millimeters serving as the validity criterion, taking into account the distance between the eyes and the tracker. Sessions with a data loss exceeding 10% were subject to exclusion and necessitated recalibration (Rudich-Strassler et al., 2022). Following calibration, participants were shown a news article on the monitor and asked to read it, with the eye-tracker monitoring their gaze throughout the reading process. It took participants approximately 6 minutes to read the news article while wearing the Tobii Pro Spark eye-tracker.

3. Data Comprehension Measurement: After reading the article on the desktop computer, participants answered questions in the Qualtrics survey pertaining to the data included in the news article using the tablet. This design allowed participants to return to the article if needed to obtain the information required to answer the questions, with eye movement during information retrieval recorded. The comprehension section took approximately 5 minutes and the survey concluded with participants submitting their responses.

To align Qualtrics data with eye-tracking data, each participant is assigned a unique Research ID in the format YYYY-MM-DD-Number, indicating their sequence of participation on that day.

Participants' demographic information is cataloged in the Tobii Pro Lab system, associated with their unique Research IDs to ensure data confidentiality while facilitating extensive analysis

through data visualization techniques. These Research IDs serve as anonymized identifiers for each participant's dataset, maintaining privacy and integrity of the research process.

3.22 Participant Recruitment

After receiving the ethics approval from Research Ethics Boards (REB) of University of Ottawa, the study's recruitment process utilized the Integrated System of Participation in Research (ISPR) Student Pool for its sampling process, with recruitment conducted directly through the portal. Managed by the School of Psychology, the ISPR offers students enrolled in participating courses in communication, psychology, linguistics, education, and business the opportunity to access and select from current research projects of interest. This study, titled "Eye-tracking the Data Consumption Patterns in Reading News Articles," was advertised within the ISPR system and offered participants one credit for their involvement, equivalent to one percent of their course grade (ISPR Student Pool, 2024). Participants were assured of receiving the full credit for their participation, irrespective of whether they completed the entire experiment or whether their data was ultimately utilized for analysis. This credit policy was implemented to encourage participation while upholding ethical standards of voluntary and informed consent.

Each experiment session was allocated a duration of 20 minutes, with an additional 20-minute buffer between sessions to provide flexibility for participants and researchers in managing the logistics of the experiment. Once the researcher created the study listing which included the necessary details and ethics approval documentation, the students were able to autonomously choose their preferred time slots and register for the in-person experiment.

3.23 Participants

The current study included 46 participants, with a gender distribution of 65.2% female ($N = 30$) and 34.8% male ($N = 16$). Participant ages varied from 18 to 39 years, with a mean age of 20.8 years ($SD = 1.66$). The vast majority, 97.8% ($N = 45$), were undergraduate students, while a single participant (2.2%) was pursuing graduate studies, reflecting a predominantly undergraduate cohort.

Linguistically, the participants presented a spectrum of English language backgrounds. A majority, 60.9%, identified English as their first language. A further 10.9% came from multilingual households with English usage, 23.9% had acquired English language proficiency during elementary or high school education, and 4.3% were comfortable with reading and speaking English, despite not having formal English education.

Regarding parental education, the participants' backgrounds were varied: 45.7% had at least one parent with an advanced degree, 34.8% had parents with undergraduate degrees, 10.9% with some form of college diploma, and 4.3% had parents with some college education but no diploma or degree. A minority, 4.3%, reported parents with a high school or GED level of education.

Table 1: Sociodemographic Characteristics of Participants

Demographics:	<i>N</i>	%
Gender		
Female	30	65.2%

Male	16	34.8%
------	----	-------

Study Level

Undergraduate 1st year	31	67.4%
------------------------	----	-------

Undergraduate 2nd year	6	13.0%
------------------------	---	-------

Undergraduate 3rd year	6	13.0%
------------------------	---	-------

Undergraduate 4th year	2	4.3%
------------------------	---	------

Other	1	2.2%
-------	---	------

English As First Language

Yes, English is my first language	28	60.9%
-----------------------------------	----	-------

No, but I come from a multi-lingual home where English was spoken.	5	10.9%
--	---	-------

No, but I learned English in elementary/high school.	8	17.4%
--	---	-------

No, but I attended an English-speaking high school.	3	6.5%
---	---	------

No, but I am comfortable reading and answering questions in English.	2	4.3%
--	---	------

Parent Education

Advanced degree	21	45.7%
-----------------	----	-------

Undergraduate degree	16	34.8%
----------------------	----	-------

College diploma	5	10.9%
-----------------	---	-------

Some college but no diploma or degree	2	4.3%
---------------------------------------	---	------

High school/GED	2	4.3%
-----------------	---	------

Age

Below and equal to 20	31	67.4%
-----------------------	----	-------

21 to 25	9	19.6%
----------	---	-------

Above 25 and others	6	13.0%
---------------------	---	-------

Note. N = 46. Participants were on average 20.8 years old (SD = 1.66).

3.24 News Article Selection

The CBC article, "How Student Loans Keep Some People Trapped in Debt," authored by Aloysius Wong and published on June 18, 2023, was selected to be the sample article. Spanning 1,308 words, this article discussed the complexities of student debt in Canada, spotlighting the hurdles faced by individuals and the broader societal implications, including issues of post-graduate equity and financial distress. Given the average reading speeds of adults—ranging between 200 to 250 words per minute (wpm) for English prose—the expected reading time was approximately 6 minutes.

The choice of this article was based on two primary considerations: (1) Data Richness: This article presented 6 data points that were presented using a blend of data categories, such as statistics, personal narratives, policy critiques, and historical perspectives. This article's diverse content offers a basis for cross-comparing many of the 15 data categories identified as independent variables in this study's eye-tracking analysis. (2) Participant Relevance: The subject of student loans was anticipated to be highly relatable for a broad audience, especially younger individuals or those connected to academic environments. Presenting participants with content that mirrors their potential personal experiences or apprehensions is likely to invoke deeper, more reflective reading.

3.25 Measures

Demographics and Disposition

Participants were surveyed on a variety of demographic attributes, including age, gender, educational background, program of study, English language proficiency, and parental education level. Gender was coded as Male = 0 and Female = 1. English proficiency was rated on a scale from 1 (comfortable with English, not a native speaker) to 5 (native English speaker), with higher scores indicating greater fluency. Parental education was categorized on a scale from 1 (High school/GED) to 5 (Advanced degree), where higher scores denote greater educational attainment.

Disposition:

Participants answered questions regarding their media interaction and mathematical background, along with their career aspirations in statistics, to measure their quantitative literacy and engagement.

Personality: Personality traits were assessed using the Big Five Inventory (BFI), conceived by John & Srivastava (1999), evaluates individuals across the Big Five personality dimensions: Extraversion vs. Introversion, Agreeableness vs. Antagonism, Conscientiousness vs. Lack of Direction, Neuroticism vs. Emotional Stability, and Openness vs. Closedness to Experience. This research focuses solely on the first dimension in each pair for assessing observer personality perceptions (John & Srivastava, 1999) to explore the potential associations with comprehension level. In this study, due to time constraints on assessment, the original 44-item BFI has been condensed into an ultra-short 10-item version (Rammstedt & John, 2007) and

assesses each personality trait with just 2 items. Participants answered 10 items, each rated on a 5-point Likert scale, which ranges from 1 ("Does not describe me at all") to 5 ("Totally describes me"). These items evaluate five personality dimensions: Extraversion, Neuroticism, Agreeableness, Conscientiousness, and Openness, with each trait represented by two items—for example, "I see myself as someone who is outgoing, sociable" and its reverse, "I see myself as someone who is outgoing, reserved." Participants demonstrated a moderate average score for three traits: Extraversion ($M = 6.50$, $SD = 2.08$), Neuroticism ($M = 6.28$, $SD = 2.00$), and Agreeableness ($M = 6.98$, $SD = 1.45$), along with relatively higher averages for Conscientiousness ($M = 7.54$, $SD = 1.60$) and Openness ($M = 7.50$, $SD = 1.57$).

Motivation: The Global Motivation Scale, as established by Guay, Blais, Vallerand, and Pelletier (1999), was utilized to evaluate participants' overall motivational orientations in line with Self-Determination Theory. Each participant responded to six questions on a 5-point Likert scale, ranging from 1 ("Does not describe me at all") to 5 ("Totally describes me"). These questions explored reasons for engaging in activities, such as, "Because I would feel bad if I do not do them," and corresponded to the various types of behavior regulation defined by the theory: intrinsic (enjoyment-driven), integrated (identity-driven), identified (value-driven), introjected (obligation-driven), external (outcome-driven), and amotivated (compulsion-driven). Scoring utilized the Self-Determination Index (SDI), as per the formula from Philippe and Vallerand (2008): $SDI = [(3 * \text{Intrinsic}) + (2 * \text{Integrated}) + \text{Identified}] - [\text{Introjected} + (2 * \text{External}) + (3 * \text{Amotivated})]$. The findings indicated robust self-determined motivation among participants, with a mean score of 6.87 and a standard deviation of 4.87. This index provides

scores ranging from -3 to 18, where higher scores reflect stronger self-determined motivational orientations.

Thought Process: Participants evaluated their thought processes using a 5-point Likert scale, from 1 ("Does not describe me at all") to 5 ("Totally describes me"). The survey questions focused on determining participants' propensity for engaging in in-depth and complex thinking. Sample statements included "I prefer complex problems to simple problems," along with its reverse, "I don't like to have to do a lot of thinking." The results indicated a strong inclination towards critical thinking and a preference for challenging cognitive tasks, with participants scoring an average of 18.07 on a scale that ranged from 8 to 25, and a standard deviation of 4.18.

News Media Knowledge: These questions evaluate a foundational comprehension of the Canadian press, media production, and consumption through six items designed to test objective knowledge of news media acumen. These questions cover topics such as the ownership of media outlets in Canada, prerequisites for becoming a news reporter, financial models of news organizations, and distinctions in news reporting between platforms like Google News and CBC News, thereby assessing literacy regarding news production. This binary scoring system used 0 for incorrect answers and 1 for correct answers of six items. The questions explored aspects of the production process and profit structure of Canadian news media. For example, one question asked, "Writing a press release would most typically be the job of:" with options such as "A) A reporter for CBC News, B) A spokesperson for Coca-Cola, C) A lawyer for Google, D) A producer for CTV's evening news broadcast, E) I don't know," where only one response was

correct. The binary scores were summed and the mean score was 2.39 ($SD = 1.45$) with scores ranging from 0 to 5, indicating a limited knowledge base regarding the media industry landscape.

News Media Attitude: Participants expressed their attitudes towards Canadian news media through a scale adapted from Primack et al. (2006)'s smoking media literacy scale, encompassing six items. The scale ranged from 1 (“I don’t know”) to 5 - positive responses, such as “The amount of racial/ethnic minority coverage in the news accurately reflects the proportion of minorities in the Canadian population.” or “People who watch a lot of television news often tend to think the world is just as violent and dangerous as it actually is.” Participants’ scores were summed across the 6 items and the results showed a generally high and consistent level of positive attitudes towards the media, with a mean score of 22.18 ($SD = 2.48$) and scores ranging from 16 to 26, indicating an overall favorable perception.

Media Experience, Confidence, and Consumption: The News Media Literacy Scale (Primack et al., 2006; Arke and Primack, 2009; Ashley et al., 2013) comprehensively assessed aspects of news media engagement, including experience, confidence, and consumption, through a total of 23 items.

Media Experience: To determine an overall picture of each participant's engagement with the media landscape, participants were asked to indicate if they had completed any of the following: media-related courses in high school (E1), in college or university (E2), completing a post-secondary diploma or degree in journalism (E3), contributing to or working for high school media outlets (W1), college or university media outlets (W2), and local media outlets (W3). The sum of these educational and work experiences yielded a total score for overall media

experience. Self-reported media experience was low, with an average score of 0.84 ($Min = 0$, $Max = 5$, $SD = 1.14$) out of a possible 6 points, indicating limited direct engagement in media production or related activities.

Media Confidence: Participants assessed their confidence in the media's accuracy and effectiveness through 18 questions using a 7-point Likert scale, which ranged from 1 ("Strongly disagree") to 7 ("Strongly agree"). To ensure the accuracy of these measurements and reduce potential response bias, the questionnaire features both positively and negatively framed statements, alongside reverse coding. Example statements included: "I don't see what news does for me," "I am in control of the information I get from the news media," and the reverse-coded "I don't think the news media can be trusted." These questions measured participants' confidence and motivation in media engagement. The participants' scores were summed across the 18 questions and Generally, participants exhibited a high level of confidence, with a mean score of 71.18 ($SD = 5.98$) with scores ranging from 59 to 82, indicating a limited range with minimal variance.

Media Consumption: The level of media consumption was measured using four questions on a 5-point scale, from 1 ("None") to 5 ("More than 2 hours"). Questions asked participants about the typical amount of time they spent engaging with different media formats daily, including online news, daily print newspapers, television news programs, and radio news programs. Participants reported an average media consumption score of 3.09 ($Min = 0$, $Max = 7$, $SD = 1.71$), suggesting a moderate amount of time dedicated to media consumption.

Subjective Numeracy: The Subjective Numeracy Scale, developed by Fagerlin et al. (2007), consists of eight items that assess participants' self-perceived numerical competence

using a 5-point Likert scale, with responses ranging from 1 (Poor) to 5 (Excellent). This scale is divided into three sections: (1) The first set of four statements includes questions such as “How good are you at calculating a 15% tip?” where participants rate their ability from 1 (Poor) to 5 (Excellent). (2) The second set of two statements asks about the usefulness of numerical data in everyday contexts, for example, “When reading the newspaper, how helpful do you find tables and graphs that are part of a story?” with scoring options from 1 (Not helpful at all) to 5 (Extremely helpful). (3) The third set of two statements explores preferences for receiving probabilistic information, such as “When people tell you the chance of something happening, do you prefer that they use words ('it rarely happens') or numbers ('there's a 1% chance')?” with responses ranging from 1 (All words) to 5 (All numbers). Overall, participants demonstrated a high level of comfort with numbers, with an average score of 28.00 ($SD = 5.39$), with scores ranging from 15 to 40. Higher scores indicate greater subjective numeracy.

Numeracy Scale: A revised version of the Objective Numeracy Scale (ONS), introduced by Nadratowski et al. (2022), assesses participants' numeracy through six mathematical questions with binary responses (0 for incorrect, 1 for correct). Each question targets a specific mathematical concept: (1) Conditional Probability, (2) Probability, (3) Percentile Operations, (4) Basic Arithmetic, (5) Exponential Function, and (6) Bayes' Theorem (Joyce, 2021). Participants selected the correct answer using a multiple-choice format in order to streamline the answering process. Correct responses were summed, and participants generally demonstrated a high level of numeracy with an average score of 3.46 ($Min = 1, Max = 6, SD = 1.81$).

Table 2: Summary of Disposition Variables

Disposition Variables	Scale	N Items
BFI-Extraversion	1-5 ^A	2
BFI-Agreeableness	1-5 ^A	2
BFI-Conscientiousness	1-5 ^A	2
BFI-Neuroticism	1-5 ^A	2
BFI-Openness	1-5 ^A	2
Thought Process	1-5 ^A	5
Self-Determination Index	1-5 ^A	6
News Media Knowledge	0,1 ^B	6
News Media Attitude	1-4/5 ^C	6
Media Experience	0-12 ^D	1
Media Confidence	1-7 ^E	18
Media Consumption	1-5 ^F	4
Math Course	0-10 ^G	4
Math Grade	0-100 ^H	1
Stats Career	1-4 ^I	1
Subjective Numeracy	1-5 ^J	8
Numeracy	0,1 ^B	6

Note: A: Likert Scale (Does not describe me at all - Totally describes me), B: Dichotomy Scale (0 = Wrong, 1 = Right), C: Likert Scale (I don't know - Positive statements (e.g., "Accurately reflects the proportion of minorities in the Canadian population.")), D: Accumulative Multiple Choices (0 = None selected, 12 = E1 + E2 + E3 + W1 + W2 + W3), E: Likert Scale (Strongly disagree - Strongly agree), F: Likert Scale (More than 2 hours - None), G and H: Scroll bar, I: Likert Scale (Not at all - A lot), J: Likert Scale (Poor/Not helpful at all/All words - Excellent/Extremely helpful/All numbers).

Eye-Tracking of News Article

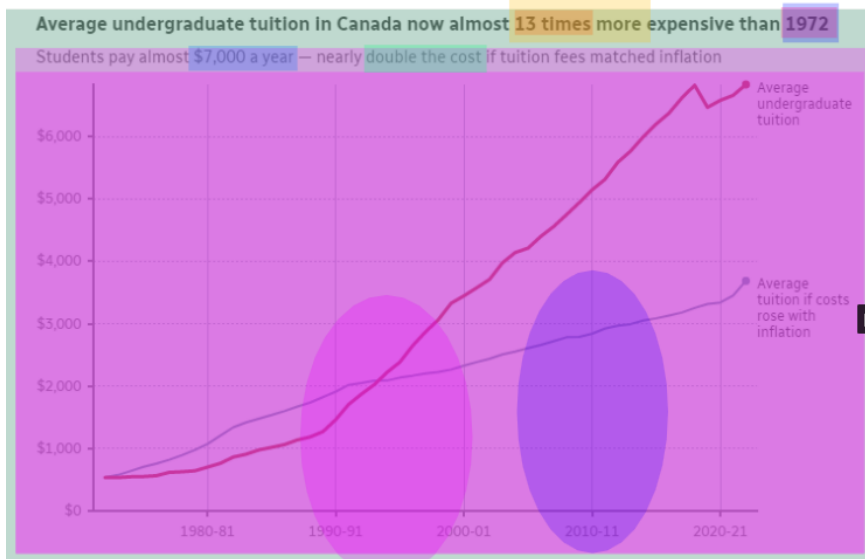
All data in the news article, regardless of presentation format, were coded as areas of Interest (AOIs) in Tobii Pro. AOIs serve as markers within an image on the screen so that the eye-tracking software can determine when participants looked in the area versus when they did not. By coding all data instances as an AOI, it allowed us to determine whether participants looked at the data and for how long, as well as make comparisons between the different AOIs.

Figure 4: Example of AOIs in News Article

High costs of higher education

Shaker says pressure to make loan payments can also trap many former students in a cycle of precarity as they take on "whatever job[s] they can" to make ends meet. Beyond that, she stresses that the increasing costs of higher education excludes those whose families can't or don't support them — and those who have children or other dependants of their own.

Average undergraduate tuition fees for full-time domestic students have risen from \$534 in 1972-73 to \$6,834 in 2022-23. These figures are almost double what they would be if tuition rates grew proportionally with inflation.



"We need a much more honest understanding of who is going to post secondary," said Shaker. "Otherwise we're leaving out swaths of students who don't fit this template and we're

Note: Areas of interest (AOIs) displaying various data types and presentations were circled in different colors.

The eye-tracking data collected for the news study includes several key metrics related to how participants read and interact with the content:

1. **Total Fixation Time on the Article:** This measures the overall time participants spent fixating on the entire article.
2. **Fixation Time on Data Elements:** This records the time spent fixating on specific data elements within the article.
3. **Fixation Time on Three Data Presentations:** This metric assesses the fixation duration on three distinct types of data presentations included in the article.
4. **Fixation Time on Five Data Types:** This evaluates how long participants fixate on five different data types featured in the article.

For the purposes of the present research, a number of variables were calculated using the AOIs. First, participants' fixation times on the AOIs were recorded, and the proportion of time dedicated to each AOI relative to the total duration of article engagement was computed to yield a ratio—a metric indicative of participant engagement levels. Ratios were calculated since individual differences in reading speed and total reading time could potentially confound the analysis. These regions were coded according to whether they contained data content (DN: Data or Not), and fell into specific data categories (DP: Data Presentation, DT: Data Type):

1. **Ratio of Data to Total Reading Time (Ratio-Data/All):** This ratio measures the proportion of time spent fixating on data elements relative to the total time spent reading the article.
2. **Ratio of Data Presentations to Total Reading Time (Ratio-DP-Raw/Narrative/Visualization):** This ratio measures the time spent viewing three types of data presentations—narrative (data presented within narrative sentences), raw (isolated numbers), and visualizations (graphics, maps, and charts)—as a proportion of the total reading time of the article.
3. **Ratio of Data Types to Total Reading Time (Ratio-DT-Absolute/Contextualizers/Proportion/Relationship/Statistical Statement):** This ratio assesses the time spent on five specific data types—absolute numbers (raw data without context), contextualizers (details like age, date, gender, socio-economic status), proportions and percentages, relationships and comparisons (such as correlation and comparative models), and statistical statements (including averages, means, and measures of central tendency)—relative to the total reading time of the article.

By adopting this ratio-centric approach, the research aims to unveil the nuances of engagement with data within news articles, circumventing the variability introduced by individual reading dynamics. The application of AOI ratios refines the analysis, isolating engagement levels from the variability inherent in personal reading styles and external disruptions. Such an approach offers a sturdy analytical base for exploring the intricacies of how readers engage with data-rich content.

Data Comprehension

Data Comprehension Questions:

Data comprehension was assessed through six individual question scores (DC1 to DC6) and an overall DC Total Score. This scoring system was used to evaluate variations in performance on single items and aggregate comprehension. Each question was scored dichotomously, with correct answers assigned a value of 1 and incorrect answers a value of 0, providing a potential score range from 0 to 6.

Table 3: Data Comprehension Questions Measurement

Number #:	Question:	Data Presentation/Type Measured:	Skill Tested:
DC1	“What year did the actual average undergraduate tuition surpass the average tuition if costs rose with inflation?”	Data Presentation: Visualization Data Type: Relationship	Graph reading
DC2	“How would you figure out the average individual amount that Canadians (with outstanding student loans) owed in July 2022?”	Data Presentation: Narrative Data Type: Statistical Statement	Information retrieval Basic arithmetic
DC3	“Of those Canadians with outstanding student loans, what was the average individual amount owed in July 2022?”	Data Presentation: Narrative Data Type: Proportion	Calculation
DC4	“How much would the average tuition have been in 2010 - 11 in Canada, if costs had risen with inflation?”	Data Presentation: Visualization Data Type: Relationship	Graph reading Information filtering
DC5	“If a student began paying off a \$10,000 loan in 2018, how much interest would they have paid in the first year?”	Data Presentation: Narrative Data Type: Proportion	Calculation

DC6	“What percentage of those who pursued professional programs took on bank loans or lines of credit (according to the 2020 Statistic Canada Report)?”	Data Presentation: Narrative Data Type: Statistical Statement	Information retrieval
-----	---	--	--------------------------

Participants had the option to refer back to the article to locate correct answers. Responses were meticulously evaluated and coded, employing a flexible approach to accuracy to accommodate a variety of potentially correct responses. Participants generally demonstrated a moderate level of comprehension, with an average DC Total Score of 2.96 ($Min = 0$, $Max = 6$, $SD = 1.63$).

Perceived Difficulty:

Following each data comprehension question, participants were asked to report their perceived difficulty in answering the question using a scale from 1 (Very Easy) to 5 (Very Difficult). This provided insights into the subjective challenge of each question (DC1 - Difficulty Score through DC6 - Difficulty Score), as well as the overall perceived difficulty level (DC Difficulty Total Score). The total allowed for a potential range from 6 (all questions rated as Very Easy) to 30 (all questions rated as Very Difficult). The aggregate DC Difficulty Total Score averaged 17.15 ($Min = 8$, $Max = 27$, $SD = 3.88$), indicating a diverse range of perceived difficulty across participants.

3.26 Analytical Plan

The analytical approach comprised two stages: preliminary analysis and model building. In the preliminary phase, the robustness of the dataset was evaluated, with a focus on identifying any

significant changes in independent and dependent variables concerning demographic factors. This step ensures the data's suitability for more complex analyses.

Subsequently, the model building phase involved constructing linear regression models to examine the relationships between variables that correlated significantly. This stage was instrumental in addressing the research questions, as it allowed for a deeper exploration of the directional and magnitude influence of various predictors on the outcomes of interest.

Initially, the data underwent a cleaning and screening process to remove any impossible values and outliers. Subsequently, descriptive statistics were computed for all variables relevant to the study. The analytical plan was then structured to address several specific research questions:

Research Question 1: *“How do different data presentations or types impact audience attention and engagement with data presented in news articles?”*

The analysis began with an examination of descriptive statistics to determine the average time and range of audience engagement (measured in seconds) with different data in the article. This step also aimed to identify any discernible patterns in how different data presentations or types influenced audience attention.

Research Questions 2.1 and 2.2: *“Does a relationship exist between the level of audience comprehension and the different data presentations or types featured in print news stories?”* and *“Does a relationship exist between the level of audience perceived difficulty and the different data presentations or types featured in print news stories?”*

The study utilized correlation analysis to explore the relationships between variables. This included examining whether there were significant associations between the time audiences spent on various data presentations or types and their levels of comprehension and perceived difficulty.

Research Question 3: *“Is there a connection between the levels of audience comprehension and their dispositional factors?”*

Statistical tests such as independent t-tests and repeated measures ANOVA were used to determine any statistically significant connections between audience comprehension levels and demographic variables.

Additionally, the study sought to develop predictive models to further the research. This involved using multilinear regression analysis and path analysis to forecast how various factors might influence audience engagement and comprehension.

Clarification of Terms:

To streamline the analysis, all variables were organized into different blocks, each designated with specific names. This classification aids in understanding and efficiently processing the data, ensuring a structured approach to answering the research questions.

Table 4: Clarification of Terms: IV, DV#1, and DV#2

Demographics:	Independent Variables (IV):	Dependent Variables#1 (DV#1):	Dependent Variables#2 (DV#2):
---------------	--------------------------------	-------------------------------------	----------------------------------

1. Gender	1. BFI-Extraversion	1. Ratio of Data to	1. DC1 Score
2. English as	2. BFI-	Total Reading Time	2. DC2 Score
First	Agreeableness	(Ratio-Data/All)	3. DC3 Score
Language	3. BFI-	2. Ratio-DP-	4. DC4 Score
3. Parent	Conscientiousness	Narrative	5. DC5 Score
Education	4. BFI-Neuroticism	3. Ratio-DP-Raw	6. DC6 Score
	5. BFI-Openness	4. Ratio-DP-	7. DC Total
	6. Thought Process	Visualization	Score
	7. Self-	5. Ratio-DT-	8. DCDifficulty1
	Determination	Absolute	9. DCDifficulty2
	Index	6. Ratio-DT-	10. DCDifficulty3
	8. News Media	Contextualizers	11. DCDifficulty4
	Knowledge	7. Ratio-DT-	12. DCDifficulty5
	9. News Media	Proportion	13. DCDifficulty6
	Attitude	8. Ratio-DT-	14. DC Difficulty
	10. Media Experience	Relationship	Total
	11. Media	9. Ratio-DT-	
	Confidence	Statistical Statement	
	12. Media		
	Consumption		
	13. Math Course		
	14. Math Grade		
	15. Stats Career		
	16. Subjective		
	Numeracy		
	17. Numeracy		

Note: For clarity in the analysis section, a list describes the different sets of variables used. The column headers (demographics, IV, DV#1, DV#2) serve as abbreviations for the terms mentioned above.

Chapter 4 Results and Analysis

4.1 Data Cleaning and Selection

Data collected from Qualtrics survey and Tobii Eye-Trackers were analyzed using IBM SPSS® software, version 26.0. Participants' datasets were carefully examined and responses were excluded if they exhibited clear response bias (uniform answers across survey questions), were incomplete due to factors such as tardiness, equipment malfunctions (e.g., software crashes, hardware issues), or were subject to environmental disruptions such as noise. In line with best practices for identifying participants who exhibited insufficient effort responding (IER; Huang et al., 2012; Meade & Craig, 2012; Ulitzsch et al., 2024), the dataset was pruned to remove participants who self-reported a preference to exclude their data. Seven participants were removed following these procedures.

Demographic information with missing or illegible responses in self-reported fields, such as age, study year, or academic program, was coded as 'Missing Value' and omitted from subsequent analysis involving the missing value(s).

Next, the identification of outliers was twofold, encompassing both univariate outliers (anomalies within a single variable) and multivariate (anomalous scores across multiple variables). Multivariate outliers were not found. However, univariate outliers were identified in the following variables: News Media Attitude ($N = 2$), Media Experience ($N = 2$), Media Confidence ($N = 1$), and Math Course ($N = 2$). In this analysis, data points were winsorized—a statistical transformation where extreme values are replaced with the nearest non-outlying values. This technique, as described by Rivest (1994), helps to mitigate the potential for extreme

values to skew results by limiting the influence of outliers. By applying Winsorization, the distribution's tails are effectively truncated, which enhances the robustness of statistical estimates against the effects of outliers (Rivest, 1994).

For the eye-tracking data, univariate outliers were noted but not removed as outliers in fixation times are pertinent to the analysis and might yield insights, especially if minimal engagement with certain data elements is observed. These outliers may point to considerable variability in how audiences scan data, a point that could be addressed in the discussion section to highlight limitations or potential areas of significance that might become more apparent with an expanded sample size.

The sample count (N) for each variable reflects the total number of observations prior to outlier curation, ensuring that all participant samples were considered in the dataset even if certain data points were classified as outliers.

4.2 Descriptive Statistics

4.2.1 Demographics and Disposition

The descriptive statistics for all study variables calculated from the Qualtrics survey are listed in Table 5.

Table 5: Descriptive Statistics of Qualtrics Survey

Dispositions	<i>M</i>	<i>Min</i>	<i>Max</i>	<i>SD</i>
---------------------	----------	------------	------------	-----------

Personality				
BFI-Extraversion	6.5	2	10	2.08
BFI-Agreeableness	6.98	4	10	1.45
BFI-Conscientiousness	7.54	4	10	1.6
BFI-Neuroticism	6.28	2	10	2
BFI-Openness	7.5	4	10	1.57
Self-Determination Index	6.87	-3	18	4.87
Thought Process	18.07	8	25	4.18
News Media Literacy				
News Media Knowledge	2.39	0	5	1.45
News Media Attitude	22.18	16	26	2.47
Media Experience	0.84	0	5	1.14
Media Confidence	71.18	59	82	5.98
Media Consumption	3.09	0	7	3.09
Mathematical and Numerical Competence				
Math Course	4.91	1	12	2.34
Math Grade	82.07	55	100	11.53
Stats Career	2.78	1	4	1.05
Subjective Numeracy	28	15	40	5.39
Numeracy	3.46	1	6	1.81

Note. N for News Media Attitude and Media Experience is 44; N for Media Confidence is 45; N for Math Course is 44; for all other variables, N = 46.

4.22 Eye-Tracking Results of News Article

Fixation Time on Article

The average time dedicated by participants to reading the entire news article was 320.05 seconds ($Min = 105.45$, $Max = 800.83$, $SD = 120.08$), translating to roughly 5 minutes per reader, corresponding to the research design. The large standard deviation (~ 2 minutes) and the extensive span of total reading times both reflect considerable variability in how participants interacted with the article, suggesting individual disparities in engagement and content processing. These metrics establish the baseline for calculating the ratios that gauge the relative distribution of attention across different content areas within the articles.

Fixation Time on Data or Not (DN) Areas of Interest (AOI)

Ratio of Data to Total Reading Time (Ratio-Data/All): Data elements were distinguished from non-data elements within the news articles by creating Areas of Interest (AOIs) specifically for content that included data. The average fixation time on these data-specific AOIs was 34.56 seconds ($Min = 5.39$, $Max = 94.55$, $SD = 20.95$). To assess participants' engagement with data, the study calculated the ratio of fixation time on data content AOIs to the overall content. This ratio had a mean of 9.92% ($Min = 3.30\%$, $Max = 19.88\%$, $SD = 4.62\%$), suggesting that participants typically spent about one-tenth of their total reading time examining the data content within the news articles.

Fixation Time on Data Presentation (DP) Areas of Interest (AOI)

Data content was parsed into distinct presentational formats—narrative, raw data, and visualized information—to explore the variations in attention attributed to each type of data element.

Data Presentations Reading Time (DP-Narrative/Raw/Visualization): The fixation time on narrative data representations (DP - Narrative) averaged 32.77 seconds (*Min* = 12.21, *Max* = 107.72, *SD* = 19.79), indicating engagement with story-driven data. For raw data (DP - Raw), the mean fixation time was shorter, at 5.76 seconds (*Min* = 0.70, *Max* = 18.97, *SD* = 4.09), suggesting a more cursory examination. Visualized data (DP - Visualization) held viewers' attention for an average of 26.31 seconds (*Min* = 0.23, *Max* = 75.41, *SD* = 17.33), with a broad span in fixation duration indicating variable engagement with graphical data presentations.

Ratio of Data Presentations to Total Reading Time (Ratio-DP-

Narrative/Raw/Visualization): The proportion of time spent on narrative data relative to the total article reading time (Ratio - DP - Narrative) had a mean of 10.08% (*Min* = 4.06%, *Max* = 20.95%, *SD* = 3.78%). The corresponding ratio for raw data (Ratio - DP - Raw) was 1.76% (*Min* = 0.41%, *Max* = 5.24%, *SD* = 0.94%), and for data visualizations, (Ratio - DP - Visualization) it was 8.06% (*Min* = 0.14%, *Max* = 20.27%, *SD* = 4.50%). These ratios serve as indicators of the attentional allocation and subsequent engagement with the different data categories within the news articles.

Fixation Time on Data Type (DT) Areas of Interest (AOI)

The research delineated data into five types within the news articles—absolute numbers, contextualizers, proportions, relationships, and statistical statements. These classifications were tagged to specific AOIs that corresponded to the described data types.

Data Types Reading Time (DT-

Absolute/Contextualizers/Proportion/Relationship/Statistical Statement): Participants spent an average of 1.67 seconds ($Min = 0.13$, $Max = 6.77$, $SD = 1.50$) fixating on absolute numbers (DT - Absolute), reflecting engagement with discrete numerical data. Contextualizers—quantitative information used to create context, including dates and currency (DT - Contextualizers)—captured attention for an average of 14.76 seconds ($Min = 2.74$, $Max = 39.37$, $SD = 8.50$), indicating a focus on contextually embedded data. The fixation time on proportions and percentages (DT - Proportion) averaged 1.20 seconds ($Min = 0.17$, $Max = 5.96$, $SD = 1.37$). For relational and comparative data (DT - Relationship), the average fixation time was notably longer at 40.96 seconds ($Min = 6.13$, $Max = 107.65$, $SD = 23.07$), with a high standard deviation suggesting individual differences in engagement with this type of data. Statistical statements (DT - Statistical statement) had an average fixation duration of 6.97 seconds ($Min = 0.38$, $Max = 31.69$, $SD = 5.05$), with the wide range implying that these elicited varying degrees of reader involvement.

Ratio of Data Types to Total Reading Time (Ratio-DT-

Absolute/Contextualizers/Proportion/Relationship/Statistical Statement): The ratio of fixation time on absolute numbers to the total article reading time was minimal, with a mean of 0.51% ($Min = 0.07\%$, $Max = 2.09\%$, $SD = 0.43\%$). Contextualizers accounted for a mean ratio of 4.52% ($Min = 1.60\%$, $Max = 10.43\%$, $SD = 1.81\%$), while proportional and percentage data had a mean ratio of 0.29% ($Min = 0.03\%$, $Max = 0.84\%$, $SD = 0.21\%$). The average fixation time ratio for relational and comparative data was more pronounced at 12.53% ($Min = 3.69\%$, $Max = 27.71\%$, $SD = 4.85\%$), with the relatively high standard deviation indicating individual

differences in engagement levels. Lastly, statistical statements comprised a mean ratio of 2.19% ($Min = 0.11\%$, $Max = 4.81\%$, $SD = 1.04\%$), again showing variability in how participants interacted with these types of data presentations. These fixation ratios serve as robust metrics to gauge participant attention and further engagement with diverse data categories within news articles.

Table 6: Descriptive Statistics of Dependent Variables

Parameters	<i>N</i>	<i>M</i>	<i>Min</i>	<i>Max</i>	<i>SD</i>
Ratio-Data/All	44	9.92%	3.30%	19.88%	3.64%
Data Presentation					
Ratio-DP-Narrative	46	10.08%	4.06%	20.95%	3.78%
Ratio-DP-Raw	46	1.76%	0.41%	5.24%	0.94%
Ratio-DP-Visualization	46	8.06%	0.14%	20.27%	4.50%
Data Type					
Ratio-DT-Absolute	41	0.51%	0.07%	2.09%	0.43%
Ratio-DT-Contextualizers	46	4.52%	1.60%	10.43%	1.81%
Ratio-DT-Proportion	38	0.29%	0.03%	0.84%	0.21%
Ratio-DT-Relationship	46	12.53%	3.69%	27.71%	4.85%
Ratio-DT-Statistical Statement	46	2.19%	0.11%	4.81%	1.04%
DC Score					
DC1-Score	46	0.74	0	1	0.44
DC2-Score	46	0.37	0	1	0.49
DC3-Score	46	0.13	0	1	0.34
DC4-Score	46	0.74	0	1	0.44
DC5-Score	46	0.3	0	1	0.47

DC6-Score	46	0.67	0	1	0.47
DC Difficulty					
DC1-DifficultyScore	46	2.76	1	5	1.1
DC2-DifficultyScore	46	2.85	1	5	1.1
DC3-DifficultyScore	46	3.61	1	5	1.1
DC4-DifficultyScore	46	1.85	1	4	0.89
DC5-DifficultyScore	46	3.33	1	5	1.14
DC6-DifficultyScore	46	2.76	1	5	1.12
DC Total Score	46	2.96	0	5	1.63
DC Difficulty Total	46	17.15	8	27	3.88

Note: "Ratio-DP-Narrative" represents the fixation time on narrative data presentations divided by the total fixation time on the article. The same calculation applies to other similar metrics.

4.23 Data Comprehension

Data Comprehension (DC) Score

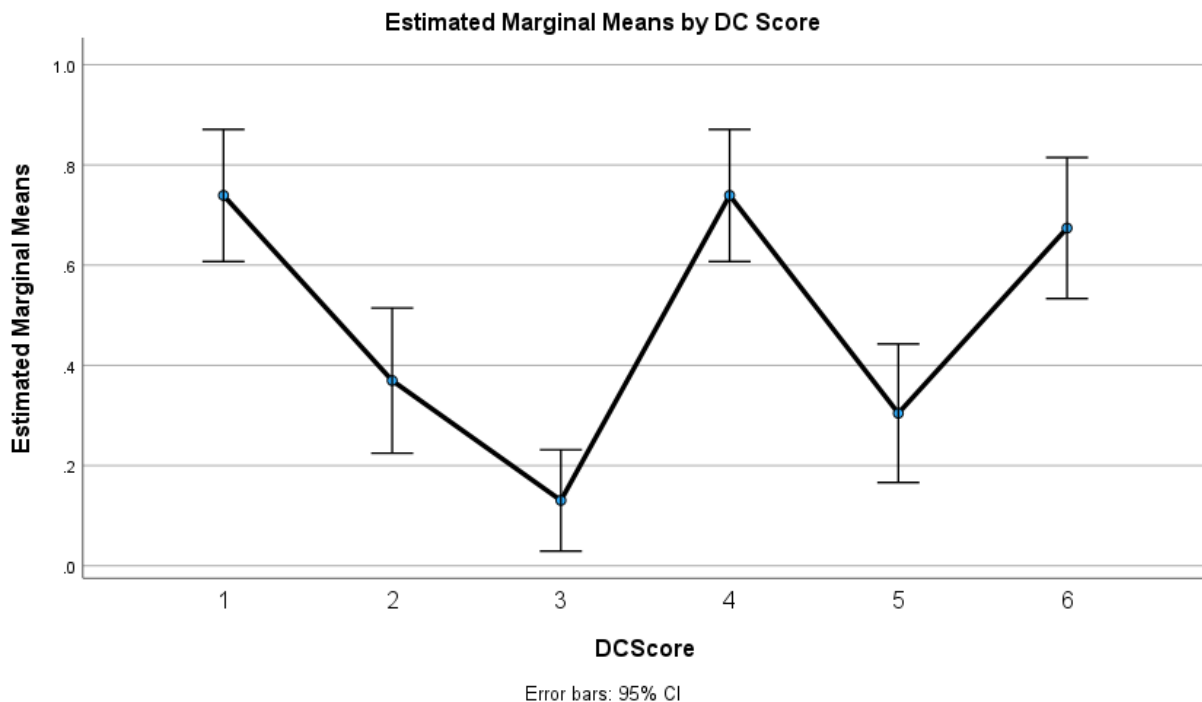
The results of the individual DC questions are visualized in Figure 4. This figure provides a graphical representation of participants' performance on each data comprehension question, showcasing the distribution and variance of scores across the six questions.

DC Score: Data Comprehension Question 1 (DC1 - Score) presented an average correct response rate of 0.74 ($Min = 0$, $Max = 1$, $SD = 0.44$), which suggests a relatively higher success rate among participants. In contrast, Question 2 (DC2 - Score), with an average of 0.37 ($Min = 0$, $Max = 1$, $SD = 0.49$), and Question 3 (DC3 - Score), with an average of 0.13 ($Min = 0$, $Max = 1$, $SD = 0.34$), indicated greater difficulty, as reflected by their lower mean scores. Questions 4 (DC4 - Score) ($Min = 0$, $Max = 1$, $SD = 0.44$) and 6 (DC6 - Score) ($Min = 0$, $Max = 1$, $SD = 0.47$)

had average scores of 0.74 and 0.67, respectively, signaling a moderate level of correct responses. Question 5 (DC5 - Score), with an average of 0.30 ($Min = 0$, $Max = 1$, $SD = 0.47$), denoted moderate complexity with a relatively lower success rate.

DC Total Score: The overall DC Total Score had an average of 2.96 ($Min = 0$, $Max = 5$, $SD = 1.63$), indicating a moderate level of comprehension among participants. This average, sitting at about 50%, suggests various potential influences on comprehension, such as the dispositional factors, data presentation, in addition to technical factors such as the constrained time in a lab setting, diverse reading habits, or environmental disturbances during the experiment.

Figure 5: Estimated Means by DC Score



Note: This figure illustrates the relationship between the DC Score (1 to 6) and the estimated marginal means of Repeated Measures ANOVA. Error bars represent a 95% confidence interval for each mean estimate.

Data Comprehension (DC) Difficulty Score

The range of perceived difficulty across the data comprehension questions is depicted in the repeated measures ANOVA shown in Figure 5.

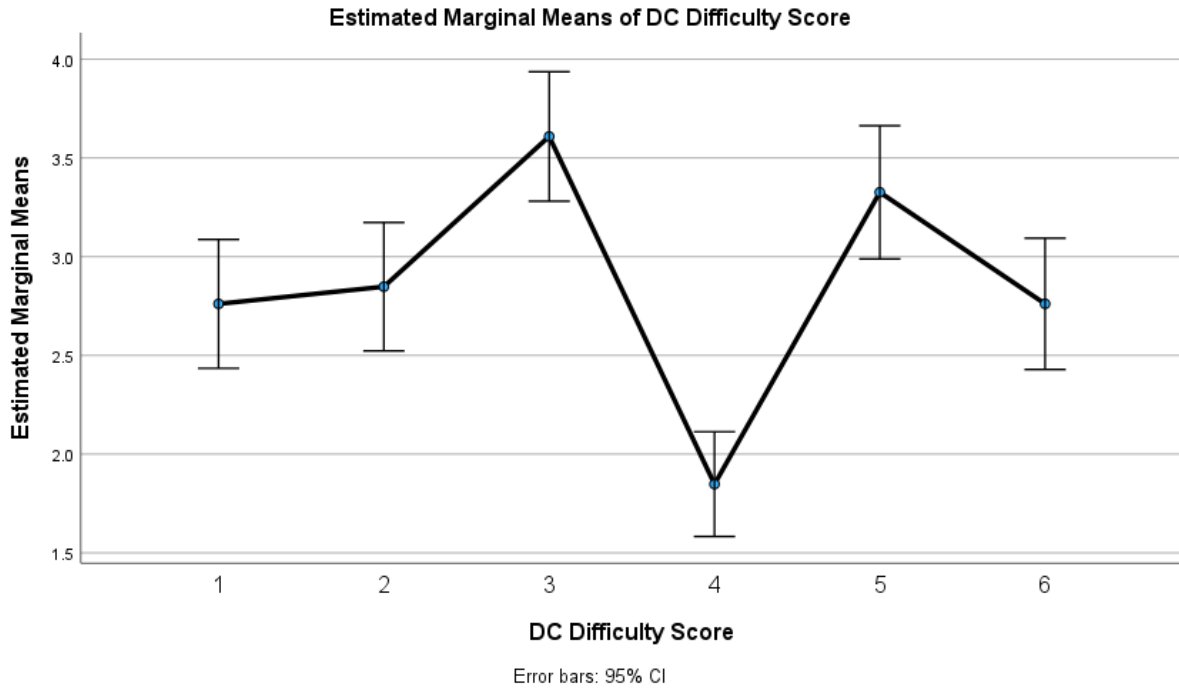
DC Difficulty Score: Data Comprehension Difficulty Question 1 (DC1 - Difficulty Score)

yielded an average perceived difficulty rating of 2.76 ($Min = 1, Max = 5, SD = 1.10$), suggesting participants found it relatively easy. Difficulty Question 2 (DC2 - Difficulty Score) had a similar average rating of 2.85 ($Min = 1, Max = 5, SD = 1.10$), while Difficulty Question 3 (DC3 - Difficulty Score) reflected a higher perceived difficulty with an average of 3.61 ($Min = 1, Max = 5, SD = 1.11$). Difficulty Question 4 (DC4 - Difficulty Score) was perceived as simpler, with an average rating of 1.85 ($Min = 1, Max = 4, SD = 0.89$), but the perceived difficulty increased for Difficulty Question 5, averaging 3.33 ($Min = 1, Max = 5, SD = 1.14$). Difficulty Question 6 then saw a decrease in perceived difficulty with an average rating of 2.67 ($Min = 1, Max = 5, SD = 1.12$).

DC Difficulty Total Score: The aggregate DC Difficulty Total Score had a mean of 17.15 ($Min = 8, Max = 27, SD = 3.88$), revealing a broad spectrum of perceived difficulty across participants.

This means that various factors could contribute to the subjective difficulty of data comprehension, including individual differences in background knowledge, cognitive processing, and question complexity.

Figure 6: Estimated Means by DC Difficulty Score



Note: This figure illustrates the average self-reported difficulty scores across six data comprehension questions and the estimated marginal means of Repeated Measures ANOVA. Difficulty was rated on a scale of 1 (Very Easy) to 5 (Very Difficult). Error bars represent a 95% confidence interval for each mean score.

4.3 Participant Background

Demographic variables such as gender, proficiency in English as a first language, academic level, and parental education were scrutinized to assess their relationship with independent variables (IVs), such as engagement with data elements in news articles, and dependent variables (DVs) like data comprehension performance. This analysis was central to exploring Research Question 3 (RQ3), which asks: "Is there a connection between audience comprehension levels and their demographic attributes?" To address this question, independent t-tests and Repeated Measures ANOVAs were employed to detect any significant differences across these categorical

demographic groups. These statistical tests helped determine whether demographic factors influenced how participants interacted with and understood the data presented in news articles.

Gender

Gender with DVs

The independent t-test results (refer to Table 23) did not reveal statistically significant differences between genders in their engagement with different data elements in news articles. Analysis demonstrated that men and women spent similar amounts of time on data elements overall, with the Ratio-Data/All yielding a $t(44) = 0.78, p = .439$. No significant gender-based disparities emerged in specific data presentation categories, such as narrative, raw data, visualization, absolute numbers, contextual information, proportional and percentile data, relational data, and statistical statements, all surpassing the $p > .05$ significance threshold.

However, gender differences were noted in certain aspects of data comprehension. The DC1-Difficulty Score revealed a significant discrepancy between genders ($t(44) = -3.13, p = .003$), with men finding question 1 easier than women. DC4 - Score ($t(44) = -2.04, p = .048$) and DC-5 Score ($t(44) = 2.17, p = .036$) both revealed slightly significant differences, with men performing better on question 5 and women on question 4.

Levene's Test for Equality of Variances indicated unequal variances between gender groups for the DC4-Score ($F = 11.03, p = .002$), DC4-Difficulty Score ($F = 4.98, p = .031$), DC5-Score ($F = 8.61, p = .005$), DC6-Difficulty Score ($F = 4.84, p = .033$), and the DC Difficulty Total ($F = 12.11, p = .001$). This suggests that individual scores varied differently for men and women (i.e.,

men may have all had similar scores on DC4-Score, while women's scores might have varied more with bigger highs and lows). Despite these unequal variances, the overall Data Comprehension Score, as indicated by both the Levene's Test ($F = 3.42, p = .071$) and the independent t-test ($t(44) = 0.318, p = .752$), did not differ significantly between genders. This suggests that individual variations in question performance and perceived difficulty did not translate into a significant difference in total comprehension scores across gender groups.

To probe the influence of gender on dependent variables, Repeated Measures ANOVAs were conducted. These analyses compared male and female participants' self-reported difficulty and the accuracy of their responses to each of the data comprehension questions (DC1 to DC6).

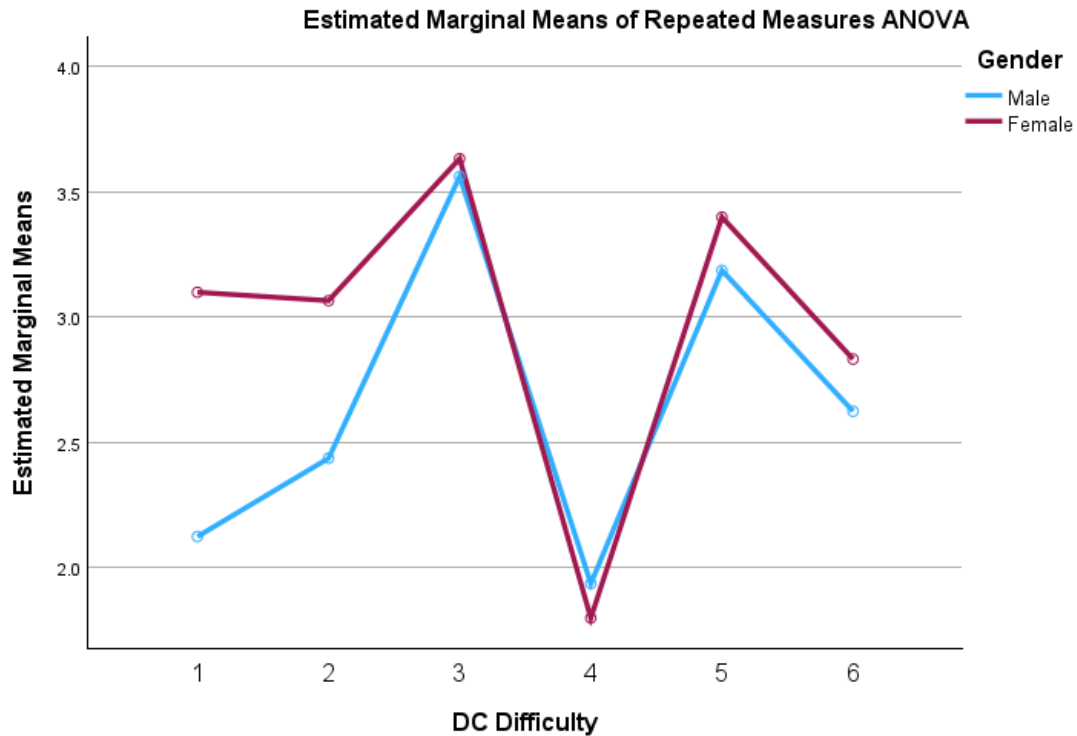
DC Difficulty (as shown in Table 7): Men ($N = 16$) reported lower average difficulty scores than women ($N = 30$) for DC1 (Men: $M = 2.13, SD = 1.15$; Women: $M = 3.10, SD = 0.92$), DC3 (Men: $M = 3.56, SD = 1.21$; Women: $M = 3.63, SD = 1.07$), and DC6 (Men: $M = 2.62, SD = 1.41$; Women: $M = 2.83, SD = 0.95$). In contrast, men indicated slightly higher difficulty for DC2 (Men: $M = 2.44, SD = 1.26$; Women: $M = 3.07, SD = 0.94$). Both genders reported comparable levels of difficulty for DC4, though women recorded a marginally lower mean score (Men: $M = 1.94, SD = 1.18$; Women: $M = 1.80, SD = 0.71$), and a similar pattern emerged for DC5, with women reporting a slightly higher mean score (Men: $M = 3.19, SD = 1.38$; Women: $M = 3.40, SD = 1.00$). The full sample ($N = 46$) exhibited the lowest mean difficulty score for DC4 ($M = 1.85, SD = 0.89$) and the highest for DC3 ($M = 3.61, SD = 1.11$), suggesting DC4 may have been perceived as less challenging overall. The relative uniformity in mean scores for DC4 and DC5

suggests a comparable level of perceived difficulty between genders for these items, while other measures display greater variation, as depicted in Figure 6.

Table 7: Repeated ANOVA of DC Difficulty Scores by Gender

DC Questions Difficulty	Gender	<i>M</i>	<i>SD</i>	<i>N</i>
DC1-Difficulty Score	Male	2.13	1.15	16
	Female	3.10	0.92	30
	Total	2.76	1.10	46
DC2-Difficulty Score	Male	2.44	1.26	16
	Female	3.07	0.94	30
	Total	2.85	1.10	46
DC3-Difficulty Score	Male	3.56	1.21	16
	Female	3.63	1.07	30
	Total	3.61	1.11	46
DC4-Difficulty Score	Male	1.94	1.18	16
	Female	1.80	0.71	30
	Total	1.85	0.89	46
DC5-Difficulty Score	Male	3.19	1.38	16
	Female	3.40	1.00	30
	Total	3.33	1.14	46
DC6-Difficulty Score	Male	2.63	1.41	16
	Female	2.83	0.95	30
	Total	2.76	1.12	46

Figure 7: Estimated Marginal Means of DC Difficulty Scores by Gender:



Note: This figure illustrates estimated marginal means for DC Difficulty Scores by gender, derived from a Repeated Measures ANOVA. The x-axis details the questions DC1-DC6, and the y-axis shows the mean difficulty scores. Red lines indicate male participants; blue lines, female participants, with error bars showing 95% confidence intervals.

Mauchly's test for sphericity was conducted to verify the assumption of equal variances across differences among the six data comprehension questions, a crucial requirement in repeated measures ANOVA to ensure equal variances of the differences between all related groups. A violation of this assumption could invalidate the F-test by increasing the Type I error rate. The test outcome was non-significant, $\chi^2(14) = 16.21, p = .302$, confirming that the assumption of sphericity had been satisfied and affirming that there are no significant differences in the data comprehension scores across various questions.

Subsequent Tests of Within-Subjects Effects revealed that the F-test for DC Difficulty was highly significant ($F(5, 220) = 17.08, p < .001$), when assuming sphericity. This result

demonstrates that significant differences exist in the perceived level of difficulty across the various questions. Further analysis through the Test of Between-Subjects Effects showed that gender, considered independently, does not have a significant impact on the perceived difficulty of data comprehension (DC Difficulty) ($F(1,44) = 2.77, p = .103$). Additionally, the interaction between DC Difficulty and Gender was not statistically significant ($F(5, 220) = 1.96, p = .086$), suggesting that gender is not a determining factor in perceived DC Difficulty.

However, univariate tests assessing male and female participants on individual questions uncovered a significant difference in DC1 ($F(1, 44) = 9.82, p = .003$). This finding, which is also visually represented in the corresponding figure, implies that while men and women generally rate the difficulty level similarly across most data comprehension questions, a notable discrepancy arises specifically with DC1. No other significant findings were observed for the remaining DC questions. Hence, when considering the repeated measures across different data comprehension questions, only DC1 showed a varied difficulty rating by gender.

DC Score (as shown in Table 8): The analysis showed that men ($N = 16$) had lower mean success scores than women ($N = 30$) for DC4 (Men: $M = 0.56, SD = 0.51$; Women: $M = 0.83, SD = 0.38$), and DC6 (Men: $M = 0.63, SD = 0.50$; Women: $M = 0.70, SD = 0.47$). Conversely, men scored marginally higher than women for DC1 (Men: $M = 0.81, SD = 0.40$; Women: $M = 0.70, SD = 0.47$) and DC2 (Men: $M = 0.44, SD = 0.51$; Women: $M = 0.33, SD = 0.48$). A significant difference in performance was observed for DC5, where men (Men: $M = 0.50, SD = 0.52$)

outperformed women (Women: $M = 0.20$, $SD = 0.41$), suggesting nuanced variations in data comprehension across genders.

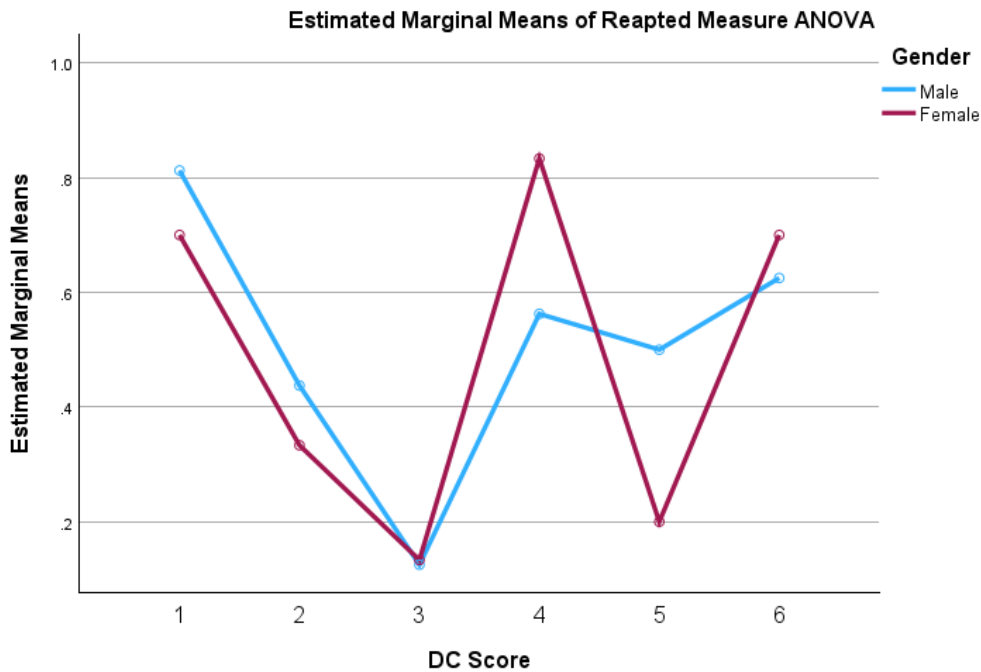
Table 8: Repeated ANOVA of DC Scores by Gender

DC Questions	Gender	<i>M</i>	<i>SD</i>	<i>N</i>
DC1-Score	Male	0.81	0.40	16
	Female	0.70	0.47	30
	Total	0.74	0.44	46
DC2-Score	Male	0.44	0.51	16
	Female	0.33	0.48	30
	Total	0.37	0.49	46
DC3-Score	Male	0.13	0.34	16
	Female	0.13	0.35	30
	Total	0.13	0.34	46
DC4 - Score	Male	0.56	0.51	16
	Female	0.83	0.38	30
	Total	0.74	0.44	46
DC5 - Score	Male	0.50	0.52	16
	Female	0.20	0.41	30
	Total	0.30	0.47	46
DC6 - Score	Male	0.63	0.50	16
	Female	0.70	0.47	30
	Total	0.67	0.47	46

These results indicate that DC3 may be the most challenging question, reflecting the lowest success rate and highest difficulty, whereas DC4 seems to be the most successfully answered

with the least difficulty reported. DC1 and DC6 show comparable levels of success and perceived difficulty. These findings correspond with the earlier analysis of individual difficulty scores, as visualized in Figure 8, illustrating the varied complexity and comprehension of data comprehension questions between male and female participants.

Figure 8: Estimated Marginal Means of Data Comprehension Scores by Gender



Note: This figure illustrates the estimated marginal means of data comprehension scores from a Repeated Measures ANOVA by gender, with the x-axis showing questions DC1 to DC6 and the y-axis the mean scores. Men are marked in red, women in blue, and error bars represent the 95% confidence intervals.

Mauchly's test for sphericity assessed the assumption of equal variances across the six individual data comprehension questions and yielded a non-significant result, $\chi^2 (14) = 21.077, p = .100$, suggesting that the assumption is met. Analysis of Within-Subjects Effects revealed highly significant F statistics for DC Score ($F (5,220) = 17.702, p < .001$), indicating significant

differences in DC scores across the questions, with multivariate tests corroborating significant performance differences between male and female participants on each question.

The Test of Between-Subject Effects revealed that the main effect of gender on DC Scores was not statistically significant ($F(1,44) = .10, p = .752$), indicating that gender alone does not have a significant impact on DC Scores.

However, the Tests of Within-Subjects Effects showed a significant interaction effect between DC Score and Gender ($F(5,220) = 2.74, p = .020$), suggesting that gender differences may influence DC Scores. Levene's Test of Equality of Error Variances revealed significant differences in variances for DC4-Score, DC5-Score, and DC6-Score when based on the mean and trimmed mean ($p = .002, .002, \text{ and } .005$, respectively). When based on the median, significance persisted for DC4-Score and DC5-Score ($p = .048$), but not for DC6-Score ($p = .615$). These results, reinforced by univariate tests, suggest a potential violation of the homogeneity of variances assumption for these scores, indicating that variance in scores may be inconsistent across genders, with significant discrepancies noted for DC4 and DC5.

In Question 4, where women outperform men, the task measures graph reading and information filtering skills. Conversely, in Question 5, where men excel over women, the task measures quick percentage calculation skills. This pattern might suggest that while women excel in filtering information and reading graphs more efficiently, men are better at performing quick mental calculations. However, these conclusions require further verification due to the limited sample size of the study.

Gender with IVs

The independent t-test results indicated statistically significant gender differences in several areas (see Table 9 in Appendix). These included: News Media Knowledge ($t(44) = 2.67, p = .011$), Math Grades ($t(44) = 2.76, p = .008$), Subjective Numeracy ($t(44) = 4.06, p < .001$), and Numeracy ($t(44) = 3.35, p = .002$). In all instances, men reported higher scores than women on average.

English As First Language

English As First Language with DVs

The One-Way ANOVA results indicated significant differences in DC3 Scores across groups based on English language proficiency ($F(4,41) = 3.04, p = .028$), suggesting variation in data comprehension scores among the groups (see Table 24 in Appendix). Tukey's HSD Post Hoc Tests (see Table 24) further revealed that those who learned English at an English-speaking high school had significantly higher DC3 Scores compared to participants from multi-lingual homes where English was spoken (Mean Difference = 0.67, $p = .043$, Standard Error = 0.23).

Additionally, a significant difference was noted when comparing the English-learnt-in-high-school group to native English speakers, with the former group again exhibiting higher scores (Mean Difference = 0.60, $p = .025$, Standard Error = 0.19). These findings highlight the potential impact of the context in which English language skills are acquired on the comprehension of data presented in English.

These results suggest that participants who learned English in high school outperformed both multilingual individuals and native English speakers on question 3, which involves information retrieval and basic arithmetic skills. Notably, this question also requires respondents to correctly answer the previous question to derive the correct response. This unexpected finding underscores the complex relationship between the context of language acquisition and data comprehension skills. It implies that individuals more proficient in English, treating it nearly as a first language, may be better equipped to correctly answer questions presented in English, highlighting the complex interplay between language acquisition context and data comprehension proficiency.

English As First Language with IVs

The analysis revealed a notable variation in the 'Statistics Career' variable across English language proficiency groups, yielding an $F(4,41) = 2.97, p = .031$, suggesting an association between English language background and participants' interest in pursuing a statistical career (refer to Table 25 in Appendix). Despite this significant ANOVA result, the subsequent Tukey HSD Post Hoc Test did not identify specific group differences that reached statistical significance (refer to Table 25). Thus, there are no significant differences in perspectives on the use of statistics in future careers based on English language proficiency. This may suggest that, regardless of language skills, participants anticipate a similar degree of statistical usage in their careers, supporting the notion that data functions as a universal language.

Parental Education Level

Parental Education Level with DVs

The one-way ANOVA results (refer to Table 26 in Appendix) identified significant differences in the Ratio of Time Spent on Absolute Data to Total Reading Time (Ratio-DT-Absolute), with an $F(4,36) = 3.46, p = 0.017$. The Tukey HSD post-hoc analysis (see Table 26) highlighted those participants with parents holding a college diploma spent a significantly greater proportion of their reading time on absolute data compared to those whose parents had undergraduate degrees ($p = 0.012$, Mean Difference = 0.76%, Standard Error = 0.22%) and those with advanced degrees ($p = 0.009$, Mean Difference = 0.76%, Standard Error = 0.21%).

A notable variance was also observed in the DC1-Score ($F(4, 41) = 2.88, p = 0.034$). The post-hoc comparison revealed that participants with "some college education but no degree" parental background underperformed on question 1 relative to peers with parents holding undergraduate degrees ($p = 0.032$, Mean Difference = -0.94%, Standard Error = 0.31).

These findings suggest that parental education levels may affect how individuals engage with and understand data in news articles. Specifically, data comprehension question 1, which involves graph reading skills and accurately recording information from a linear graph, indicates that children of parents with higher education levels may have better training in graph reading and information retrieval. However, the limited instances of significant differences call for caution, as these results could be due to statistical artifacts rather than genuine underlying relationships. Furthermore, the conclusions drawn require further examination and verification due to the limited sample size of this study.

Parental Education Level with IVs

The analysis indicated no significant differences in the independent variables across the various categories of parental education levels (refer to Table 27 in Appendix).

4.4 Correlations

A detailed cross-comparison involving demographics, independent variables (IVs), and dependent variables (DVs) was conducted to unearth potential correlations, thereby addressing the primary research inquiries: What elements influence data comprehension outcomes? (Refer to Table 28 in Appendix). This exploratory analysis is geared toward pinpointing pivotal relationships among the variables.

In tackling Research Question 2.1 (RQ2.1): "*Does a relationship exist between the level of audience comprehension and the various data presentations or types featured in print news stories?*" and its subsequent query, RQ2.2: "*Does a relationship exist between the level of audience perceived difficulty and the different data presentations or types featured in print news stories?*" The study assigns a hierarchy to variables based on their significance and utility. The variables are ranked as follows: 1) Eye-tracking metrics, which quantify the time ratios spent on distinct data elements (DVs #1), and 2) Outcomes of data comprehension, as depicted by DC Scores and DC Difficulty Scores (DVs #2) (refer to Table 4). These elements are crucial for examining how different data presentations or types influence understanding.

Furthermore, to explore Research Question 3 (RQ3): "*Is there a connection between audience comprehension levels and their dispositional factors?*" the study extends its focus to include, in addition to DVs#1 and DVs#2, 3) Disposition metrics (IVs), and 4) Detailed demographic data. This approach steers the investigation toward understanding the multifaceted influences on data comprehension, for further prediction and model building.

Relationship Between Attention Span and Data Comprehension Performance: Dependent Variables 1 and 2 (DVs #1 & DVs #2)

Table 28 reveals significant relationships, notably between the Ratio-Data/All and the DC Difficulty Score ($r = -.35, p = .020$), indicating a moderate inverse correlation between the proportion of time spent on data elements and the perceived difficulty of data comprehension questions. This suggests that investing more time in reading data elements may reduce their perceived difficulty. Additionally, a significant relationship was observed between Ratio-DR-Visualization and the DC Difficulty Total ($r = -.29, p = .048$), suggesting that increased time spent on graphs and visualizations is associated with lower perceived difficulty. However, no significant relationships were detected between the attention spans allocated to various data elements and the overall data comprehension level (DC Total Score). This indicates that while the time ratio spent on data elements contributes to reducing perceived difficulty, it does not necessarily affect actual comprehension performance.

Relationship Between Dispositions and Data Comprehension Performance: Independent Variables (IVs) and Dependent Variable 2 (DVs #2)

Significant strong correlations were identified between the Self-Determination Index (SDI) and DC Difficulty Total ($r = -.44, p = .002$), as well as between the SDI and DC Total Score ($r = .44, p = .003$). These findings suggest that participants with higher levels of self-determination not only perceive data comprehension questions as easier but also tend to perform better on these questions. This suggests that a higher degree of self-motivation may enhance both the perception and actual performance in data comprehension.

Similarly, a strong correlation was observed between Thought Process (TP) and DC Difficulty Total ($r = -.44, p = .002$), indicating that individuals inclined towards accepting challenging and reflective thoughts and engaging in intricate thinking processes tend to perceive the questions as less difficult. However, the absence of a significant correlation between TP and DC Total Score suggests that the influence of self-reported thinking patterns may not directly impact overall comprehension performance.

Moderate correlations were also detected between News Media Knowledge (NMK) and DC Difficulty Total ($r = -.33, p = .027$), suggesting that a more profound knowledge of the news media production process and an understanding of the industry's inner workings are associated with a lower perceived difficulty of data comprehension questions. This implies that a high level of contextual knowledge about the industry may facilitate easier comprehension, though this does not necessarily translate into improved performance on DC questions.

Furthermore, moderate correlations were found between Subjective Numeracy and DC Difficulty Total ($r = -.29, p = .049$), indicating that higher self-assessed numeracy skills are associated with lower perceived difficulty. This finding points to the possibility that confidence in numerical and mathematical skills may influence participants' perceptions when faced with data comprehension questions. A similar pattern was noted between Numeracy and DC Difficulty Total ($r = -.33, p = .027$), where higher objective numeracy scores correlate with perceived ease. However, these subjective and objective numeracy measures do not significantly impact the actual comprehension performance, suggesting that while robust mathematical confidence and skills may reduce perceived difficulty, they do not necessarily enhance understanding of data in news articles.

Finally, moderate correlations were found between Statistics Career Aspiration and DC Total Score ($r = -.34, p = .0019$), indicating that aspirations towards a career in statistics are closely linked with better data comprehension scores. This suggests that individuals with a clear interest in statistics may have enhanced abilities or greater motivation in handling quantitative data effectively. This aligns with the theories proposed by Ramirez et al. (2016) and Wang et al. (2015), which suggest that the necessity of math skills and the enhancement of these skills can be influenced by exposure and motivation related to the use of math in one's career.

Relationship Between Dispositions and Attention Spans: Independent Variables (IVs) and Dependent Variable 1 (DVs #1)

As a final part in correlation analysis, an investigation into the relationships between disposition metrics and eye-tracking metrics was conducted, focusing on participants' dispositions and their

attention spans towards data elements within news articles. Notably, significant relationships emerged between individual levels of conscientiousness (BFI-Conscientiousness) and the proportional time spent on raw data (Ratio-DR-Raw) as well as on proportional and percentile data (Ratio-DT-Proportion), with correlation coefficients of $r = -.30$ ($p = .040$) and $r = -.38$ ($p = .021$), respectively. This pattern suggests that individuals characterized by high conscientiousness, known for their efficient, organized, and thorough approach (Costa Mastrascusa et al., 2023), may navigate and process specific data elements, like raw and percentile data, more swiftly. This notable correlation invites further investigation to delve into the nuances of this relationship and its implications for data comprehension strategies.

Additionally, a significant relationship was observed between the level of News Media Knowledge (NMK) and the proportional time spent on statistical statements (Ratio-DT-Statistical Statement) ($r = -.32$, $p = .030$), indicating that more extensive industry knowledge correlates with reduced reading time on statistical statements. This suggests that a rich knowledge base may facilitate faster comprehension of statistical information within news articles.

Moreover, significant relationships were discovered between the level of News Media Attitude (NMA) and the proportional time spent on narrative data (Ratio-DR-Narrative) and contextualizers (Ratio-DT-Contextualizers), with correlation coefficients of $r = .36$ ($p = .015$) and $r = .32$ ($p = .037$), respectively. These findings indicate that a more positive attitude towards the credibility and effectiveness of news media is associated with increased attention span and reading time for narrative and contextualizing data elements.

Significant relationships were also noted between Subjective Numeracy and the proportional time spent on all data elements (Ratio-Data/All) ($r = .31, p = .042$), suggesting that higher self-reported numeracy skills are linked to longer data element reading times. In contrast, this significant relationship was not observed with objective Numeracy, indicating a discrepancy between subjective and objective numeracy skills in terms of data engagement. This discrepancy highlights the difference in behavior between participants with high self-assessed numeracy and those with high objective numeracy scores, especially in reading speed and attention span. Additionally, a significant relationship was found between Subjective Numeracy and the proportional time spent on relational and comparison data (Ratio-DT-Relationship) ($r = .30, p = .042$), suggesting that higher self-reported numeracy leads to increased attention and reading time. However, no significant relationship was found between objective Numeracy and Ratio-DT-Relationship, underscoring the distinction between subjective and objective numeracy levels in reading behaviors.

4.5 Predictive Model

Having established descriptive outcomes and identified correlations between dependent and independent variables, along with the potential influence of demographics and dispositions on dependent variables, the study advanced its exploration. To address Research Question 3—"*Is there a link between audience comprehension levels and dispositional factors?*"—predictive model analyses were employed to thoroughly assess the proposed conceptual framework.

Regression

In the initial phase of predictive modeling, multiple regression analysis was conducted, structured around the four disposition blocks delineated in the variable design phase. Block 1 encompasses Personality, as indicated by the Big Five Inventory (BFI) scores. Block 2 captures the Critical Stance, reflected by the Self-Determination Index and Thought Process metrics. Block 3 comprises News Media Literacy, represented by News Media Knowledge (NMK), News Media Attitude (NMA), Media Experience, Media Confidence, and Media Consumption. Block 4 relates to Mathematical Knowledge, incorporating variables such as Math Course, Math Grade, Statistics Career, Subjective Numeracy, and Objective Numeracy. Gender was included as a covariate throughout. This exploratory step aimed to identify discernible patterns within the regression models that could warrant further detailed analysis.

Personality Prediction on DVs

Hierarchical multiple regression analysis was employed to predict DC Total Score based on personality traits from the Big Five Inventory (BFI), as reported in Table 10 (see Appendix). The first model, incorporating BFI-Extraversion, accounted for a negligible variance in DC Total Score ($R^2 = .017$), which was statistically insignificant ($F(1, 44) = 0.76, p = .387$). Introduction of BFI-Agreeableness in the second model did not significantly alter the predictive capacity ($\Delta R^2 < .01, \Delta F = 0.07, p = .800$). The inclusion of BFI-Conscientiousness in the third model similarly failed to produce a significant variance increase ($\Delta R^2 < .01, \Delta F = 0.20, p = .656$). With BFI-Neuroticism added in the fourth model, a slight upturn in explained variance was noted ($\Delta R^2 = .06$), nearing significance ($\Delta F = 2.90, p = .096$), hinting at a potential link between neuroticism and DC Total Score. The final model with BFI-Openness did raise the R^2 to .10, but this

increment was not significant ($\Delta R^2 < .01$, $\Delta F = 0.35$, $p = .556$), and the adjusted R^2 declined marginally, suggesting no substantial contribution of openness to variance in DC Total Score.

The aggregate analysis shows that individual BFI personality factors lead to minor R^2 fluctuations, yet these do not reach statistical significance, nor does the model's predictive efficacy improve substantially with their inclusion. The consistently low or negative Adjusted R^2 values across models indicate that the personality traits measured by the BFI are not significant predictors of overall data comprehension level.

When applying the same predictors to DC Difficulty Total, as detailed in Table 11, all five models were statistically insignificant ($\Delta F (1, 44) < 1.07$, $p > .308$), demonstrating that the Big Five Inventory (BFI) traits are not significant predictors of the overall self-reported difficulty scale.

SDI and Thought Process Prediction on DVs

The hierarchical regression analysis investigated the predictive capacity of two models for the DC Total Score (see Table 12 in Appendix). Model 1, incorporating the Self-determination Index (SDI) alone, accounted for a significant proportion of the variance in DC Total Score, with an R^2 of .189. This indicates that approximately 18.9% of the variance in DC Total Score is attributable to the SDI, or participants' general motivation. The model is statistically robust ($\Delta F (1, 44) = 10.27$, $p = .003$). The introduction of Thought Process (TP) in Model 2 yielded a marginal increase in R^2 to .191; however, this increment was not statistically significant ($\Delta R^2 < .01$, $\Delta F = 0.09$, $p = .769$). The Adjusted R^2 for Model 2 was .153, showing a slight decrease compared to

Model 1's Adjusted R^2 of .171, suggesting that TP does not enhance the model's predictive accuracy for DC Total Score. The standard error of estimate rose from 1.48 in Model 1 to 1.50 in Model 2, further suggesting that TP's inclusion does not improve predictive precision.

Overall, while the SDI is a significant predictor of DC Total Score, the addition of TP does not significantly increase the model's predictive strength. The findings reflect prior correlations identified, with SDI having a significant correlation with DC Score and TP showing no correlation. However, the significant correlation between SDI and TP ($r = .38, p = .009$) necessitates a careful evaluation of potential multicollinearity effects, which may call for model adjustment.

Next, two models were examined to predict DC Difficulty Total (see Table 13 in Appendix). Model 1, featuring the Self-determination Index, significantly predicts DC Difficulty Total, accounting for 19.4% of its variance ($R = .44, R^2 = .19, \text{Adjusted } R^2 = .175, \Delta F(1, 44) = 10.57, p = .002$). Introducing Thought Process as a second predictor in Model 2 increases the R^2 to .28, indicating that together, these predictors explain approximately 28.3% of the variance in DC Difficulty Total. The Adjusted R^2 also rises to .25, signifying an improved model fit with the inclusion of an additional predictor. The contribution of Thought Process is statistically significant, evidenced by a notable R^2 change of .09 ($\Delta F = 5.39, p = .025$).

This significant ΔF and the R^2 improvement confirm that Thought Process significantly enhances the model's predictive capability for DC Difficulty Total, beyond the explanation provided by the Self-determination Index alone. The reduction in the standard error of the estimate from 3.52

in Model 1 to 3.36 in Model 2 further illustrates increased prediction accuracy with the integration of Thought Process. Similar to previously mentioned multicollinearity concerns, despite the model's good fit, it requires further scrutiny and potential adjustment to address these issues adequately.

News Media Literacy Prediction on DVs

A hierarchical multiple regression analysis was conducted to ascertain the role of news media-related predictors in accounting for the variance in DC Total Score, as reported in Table 14 in Appendix.

Model 1, which included News Media Knowledge, explained a small 3.4% of the variance in DC Total Score ($R^2 = .03$), although the contribution was not statistically significant ($\Delta F(1, 44) = 1.57, p = .217$). Adding News Media Attitude in Model 2 slightly elevated the explained variance to 9.5% ($R^2 = .095$); yet the increase in R^2 did not achieve statistical significance ($\Delta R^2 = .06, \Delta F = 2.90, p = .096$). Model 3, which incorporated Media Experience, expanded the variance explained to 13.7% ($R^2 = .137$), but the increment was not statistically significant ($\Delta R^2 = .04, \Delta F = 2.04, p = .161$). Introducing Media Confidence in Model 4 resulted in a minor and non-significant increase in explained variance to 14.1% ($R^2 = .14, \Delta R^2 < .01, \Delta F = 0.19, p = .667$). Lastly, Model 5 included Media Consumption, raising the explained variance to 14.5% ($R^2 = .15$), with no significant change from Model 4 ($\Delta R^2 < .01, \Delta F = 0.19, p = .662$).

The low Adjusted R^2 values across the models, ranging from .01 to .08, suggest the limited explanatory power of these news media-related variables for DC Total Score when the number of

predictors is considered. The standard error of the estimate showed a slight decrease from Model 1 to Model 3, hinting at marginal gains in predictive accuracy with the addition of new predictors. However, Models 4 and 5, with non-significant R^2 changes, appear not to provide significant incremental predictive value. The analysis indicates that while the included news media-related variables are theoretically linked to DC Total Score, none produced a statistically significant increase in explained variance, suggesting their limited predictive utility in this context or a possibly insufficient sample size to detect their effect.

Next, five models progressively incorporated additional predictors to examine their collective predictive power for DC Difficulty Total (see Table 15 in Appendix).

Model 1, using News Media Knowledge alone, explains 10.7% of variance in DC Difficulty Total ($R^2 = .11$), achieving statistical significance ($F(1, 44) = 5.25, p = .027$). Model 2 adds News Media Attitude, slightly raising the R^2 to .15. However, this increase ($\Delta R^2 = .04$) does not significantly boost the model ($\Delta F = 2.09, p = .156$). Model 3 introduces Media Experience, with a negligible R^2 change ($< .01$) and no statistical significance ($p = .772$). Model 4, incorporating Media Confidence, sees a similarly small R^2 increase ($< .01$) without significant model improvement ($p = .759$). Model 5 includes Media Consumption, elevating the R^2 to .16, indicating 16.4% variance explained in DC Difficulty Total, with an Adjusted R^2 of .06. The transition from Model 4 to 5 remains statistically insignificant ($\Delta R^2 = .01, \Delta F = 0.58, p = .449$).

News Media Knowledge in Model 1 is the sole statistically significant predictor for DC Difficulty Total. The subsequent inclusion of News Media Attitude, Media Experience, Media Confidence, and Media Consumption fails to significantly enhance predictive capability.

Math Competence Prediction on DVs

In the hierarchical multiple regression analysis examining predictors of DC Total Score, the model's efficacy evolved with the inclusion of various predictors, as reported in Table 16 in Appendix.

Model 1, which solely featured Math Course, did not account for any variance in DC Total Score ($R^2 < .01$), evidenced by an inconsequential R^2 and Adjusted R^2 , and a non-significant change in R^2 ($\Delta F < .01, p = .962$). Adding Math Grade in Model 2 marginally increased the explained variance ($R^2 = .007$); however, this enhancement was not statistically significant ($\Delta F = .32, p = .576$), leaving the Adjusted R^2 negative and indicating a lack of significant improvement from Math Grade. Model 3, with the inclusion of Stats Career, marked a noticeable increase in explained variance ($R^2 = .12$), with a statistically significant model change ($\Delta F = 5.57, p = .023$), signifying Stats Career's substantial contribution to predicting DC Total Score. Introducing Subjective Numeracy in Model 4 raised the R^2 to .18, though this increase from the preceding model did not reach statistical significance ($\Delta F = 3.05, p = .088$), albeit nearing it. The improvement in Adjusted R^2 suggests an added explanatory benefit from Subjective Numeracy. Model 5, incorporating Numeracy, presented the highest variance explained ($R^2 = .22$), despite the non-significant incremental change with Numeracy's addition ($\Delta F = 2.04, p = .161$). The Adjusted R^2 peaked at .13, hinting at some incremental explanatory utility.

The model's ability to predict DC Total Score saw gradual improvement with the sequential addition of variables, particularly noting a significant boost with the inclusion of Stats Career. The comprehensive model, incorporating Math Course, Math Grade, Stats Career, Subjective Numeracy, and Numeracy, accounted for approximately 22.4% of the variance in DC Total Score, aligning with initial exploratory correlation analyses. However, considerations for multicollinearity are warranted due to notable correlations among variables: Math Course and Subjective Numeracy (SN) ($r = .39, p = .008$), Math Course and Numeracy (N) ($r = .32, p = .028$), Math Grade and SN ($r = .45, p = .002$), Math Grade and N ($r = .40, p = .006$), and between SN and N ($r = 0.52, p < .000$). Collinearity diagnostics underscored this concern, indicating a need for model refinement to address the potential influence of multicollinearity on the analysis outcomes.

Next, the incremental effects of adding academic and psychological math knowledge factors to predict DC Difficulty Total were examined (see Table 17 in Appendix).

Model 1 features Math Course, yielding an insignificant R^2 of below .01 ($\Delta F(1, 44) = .03, p = .872$), indicating minimal predictive power for DC Difficulty Total. Adding Math Grade in Model 2 slightly raises the R^2 to .01, but the increase lacks significance ($\Delta F = 0.35, p = .560$), and the Adjusted R^2 remains negative, showing no substantial model improvement. Model 3's inclusion of Stats Career improves R^2 to .06; however, the change is not significant ($\Delta F = 2.15, p = .150$). Subjective Numeracy in Model 4 elevates R^2 to .13, nearing significance ($\Delta F = 3.25, p = .079$) and enhancing Adjusted R^2 , hinting at a stronger link with DC Difficulty Total. Model 5

introduces Numeracy, achieving the highest R^2 of .20, suggesting it accounts for 20% of variance in DC Difficulty Total. This model shows the most notable increase ($\Delta R^2 = .07$, $\Delta F = 3.72$, $p = .061$), with Adjusted R^2 climbing to .10, indicating a 10% variance explanation when adjusting for predictors.

While incremental increases in explained variance for DC Difficulty Total are observed across models, none achieve statistical significance at the .05 level. Nonetheless, the addition of Numeracy displays a promising trend towards significance, highlighting its potential as a valuable predictor. Examination of the Excluded Variables Table, detailed in Table 18, reveals Numeracy as a consistently significant predictor of DC Difficulty Total, whereas Subjective Numeracy's significance varies across models. Neither Stats Career nor Math Grade demonstrate significant predictive power in any model. The presence of negative beta coefficients for Numeracy indicates an inverse relationship, where higher numeracy skills correlate with reduced difficulty in data comprehension tasks. This positions both Subjective Numeracy and Numeracy as key predictors for further analytical exploration.

Table 18: Result of Excluded Variables in “Math Competence as Predictor Model”

Model	Beta In	t	Sig.	Partial Correlation	Collinearity Statistics			
					Tolerance	VIF	Minimum Tolerance	
1	Math Grade	-.094 ^b	-0.587	0.560	-0.089	0.900	1.111	0.900
	Stats Career	-.215 ^b	-1.435	0.158	-0.214	0.984	1.016	0.984
	Subjective Numeracy	-.333 ^b	-2.110	0.041	-0.306	0.848	1.179	0.848
	Numeracy	-.355 ^b	-2.346	0.024	-0.337	0.899	1.113	0.899

2	Stats Career	-.222 ^c	-1.465	0.150	-0.220	0.981	1.020	0.883
	Subjective Numeracy	-.344 ^c	-2.003	0.052	-0.295	0.732	1.366	0.732
	Numeracy	-.364 ^c	-2.242	0.030	-0.327	0.799	1.252	0.799
3	Subjective Numeracy	-.311 ^d	-1.802	0.079	-0.271	0.714	1.400	0.714
	Numeracy	-.400 ^d	-2.514	0.016	-0.366	0.787	1.271	0.787
4	Numeracy	-.335 ^e	-1.928	0.061	-0.292	0.661	1.512	0.600

Note: a. Dependent Variable: DC Difficulty Total; b. Predictors in the Model: (Constant), Math Course; c. Predictors in the Model: (Constant), Math Course, Math Grade; d. Predictors in the Model: (Constant), Math Course, Math Grade, Stats Career; e. Predictors in the Model: (Constant), Math Course, Math Grade, Stats Career, Subjective Numeracy

DC Difficulty Prediction on Data Comprehension

In the final analytical step of this section, the study employed regression analysis to assess DC Difficulty Total as a predictor for DC Total Score (see Table 19 in Appendix). The R^2 value of .27 signifies that approximately 27.3% of the variance in DC Total Score is accounted for by DC Difficulty Total ($\Delta F(1,44) = 16.51, p < .001$). The Adjusted R^2 of .26, adjusting for the number of predictors, indicates a robust model fit considering the model's complexity. The highly significant p-value ($< .001$) underscores a strong relationship, highlighting those variations in DC Difficulty Total correlate with fluctuations in DC Total Score. Nonetheless, with over 70% of the variance in DC Total Score remaining unexplained by DC Difficulty Total alone, it is clear that additional factors likely contribute to the overall prediction of DC Total Score.

The findings from the regression model exploration are summarized in the following table (see Table 20 in Appendix), which categorizes different building blocks into groups and examines

two models: one with data comprehension level as the dependent variable and another with perceived difficulty as the dependent variable. The summary table clearly compares the factor loads of each dispositional factor.

Path Analysis

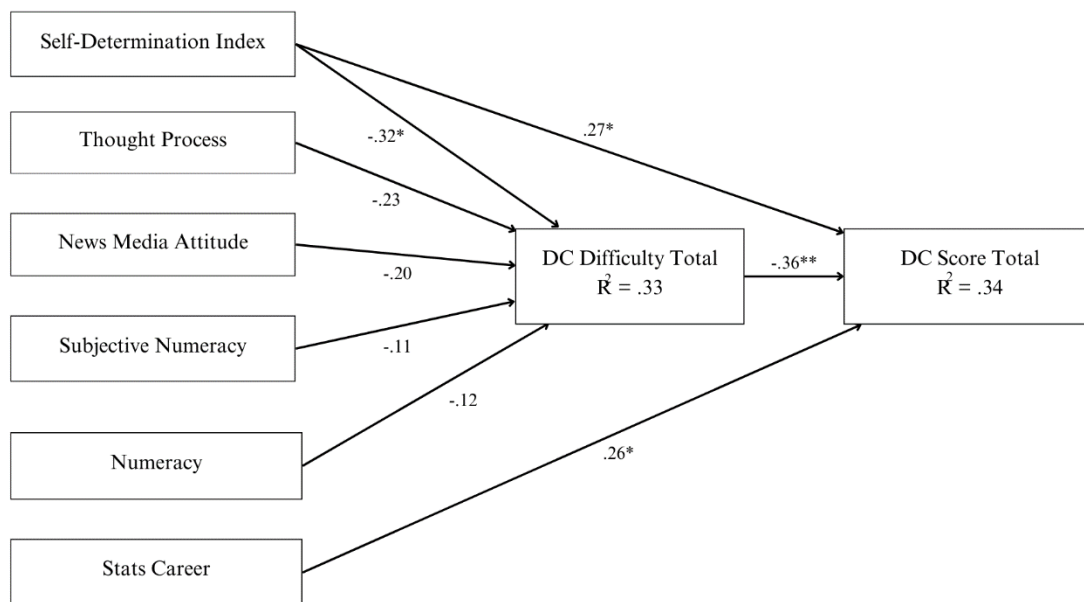
Upon establishing the correlations, the study advanced to path analysis with the objective of constructing a predictive model for data literacy, simultaneously addressing the collinearity detected in prior regression analyses. Path analysis is a statistical technique used to examine the direct and indirect relationships between variables in a hypothesized model. It allows us to specifically trace the causal connections between variables, helping to clarify which factors directly influence an outcome and which play a mediating role (Kline, 2016). The structural model testing phase entailed evaluating construct collinearity, verifying the significance of path coefficients, employing bootstrapping techniques for path analysis, and appraising the model's predictive capacity and integrity according to the standards delineated by Hair et al. (2019). The deployment of SPSS AMOS facilitated a nuanced examination of the regression weights for each independent variable and the overall model fit. This analytical process yielded a general model that demonstrated a robust ability to predict news literacy, subsequently undergoing multigroup analysis to discern the influences and moderating roles of gender, ultimately endeavoring to establish a Structural Equation Model (SEM) in harmony with Gal's (2002) conceptualization of data literacy.

The construction of the multiple regression model initiated with a preliminary investigation to identify potential predictors, revealing that while certain variables predict perceived difficulty

(DC Difficulty)—such as the Self-Determination Index (SDI), Thought Process (TP), News Media Knowledge (NMK), Subjective Numeracy (SN), and Numeracy (N)—only a select few, specifically SDI, a career in statistics (Stats Career), and DC Difficulty, emerged as predictors for the comprehension scope. These variables were chosen based on their significant correlation with the dependent variables (DC Difficulty and DC Score), ensuring the model structure adequately explains the variance in these outcomes. Following this principle, correlations lacking statistical significance were excluded from the analysis to maintain a model that strikes a balance between fit and parsimony.

The development of the general model is depicted in Figure 8. A bootstrapping procedure, applying a 5% significance threshold with 2,000 subsamples, was executed to evaluate the significance of indirect effects within mediation models.

Figure 1: Path Analysis Model



Note: Path analysis modeling the relationships between Self-Determination Index, thought process, news media knowledge, subjective numeracy, numeracy, statistics career interest, total

*DC difficulty, and total DC score. * $p < .05$. ** $p < .01$. *** $p < .001$. Numbers above the arrows represent beta weights.*

Table 21 presents the outcomes of the path analysis with standardized regression weights. The Self-determination Index emerges as a formidable predictor of DC Difficulty Total, exhibiting a moderate negative effect ($\beta = -.32$). Thought Process also shows a moderate negative influence ($\beta = -.23$) on DC Difficulty Total. Compared to the Self-determination Index, the impacts of News Media Knowledge ($\beta = -.20$) and Numeracy ($\beta = -.12$) are less pronounced. Stats Career positively correlates with DC Total Score ($\beta = .26$), signifying a moderate beneficial impact. Moreover, DC Difficulty Total inversely affects DC Total Score ($\beta = -.36$), indicating that higher difficulty corresponds to lower scores. The Self-determination Index, with a positive impact ($\beta = .27$) on DC Total Score, is pivotal in affecting both DC Difficulty Total and DC Total Score, demonstrating statistical significance ($p < .05$). Stats Career also significantly influences DC Total Score ($p < .05$), whereas other model variables exhibit negative associations with DC Difficulty Total, albeit not reaching statistical significance ($p > .05$).

In terms of model efficacy, R^2 represents the exogenous variables' cumulative impact on endogenous variables. Acceptable R^2 values, as noted by Hair et al. (2014), are 0.25 (weak), 0.50 (medium), and 0.75 (high). According to Table 21, this model accounts for 33% of the variance in total perceived difficulty and 34.1% in comprehension level, denoting a moderate predictive power.

Model Fit Indices:

- The Chi-Square (CMIN) value of 10.363 with 15 degrees of freedom ($p = .796$) indicates an acceptable model fit, with the p -value $> .05$ suggesting no significant discrepancy between the model and observed data.
- A χ^2/df ratio of .691 falls below the recommended range of 2 to 5 by Tabachnick and Fidell (2007) and Wheaton et al. (1977), potentially due to the small sample size.
- The Root Mean Square Residual (RMR) at .667 exceeds the preferred maximum of .08, implying a suboptimal fit, while the Goodness-of-Fit Index (GFI) at .947 approaches 1, indicating a good fit.
- The Adjusted Goodness-of-Fit Index (AGFI) at .873, although below the ideal .90, is deemed acceptable.
- An RMSEA of .000 with a PCLOSE of .862 exemplifies an excellent model fit, with a high likelihood that RMSEA is under .05.
- Fit indices such as the NFI (.875), TLI (1.158), and a perfect CFI (1.000) suggest acceptable to excellent model fits.
- Parsimony indices (PRATIO, PNFI, PCFI) point to a moderately parsimonious model.

In summary, despite not achieving perfection across all fit indices, the model shows commendable performance, particularly in CFI, TLI, and RMSEA, marking it as a reasonably good fit overall.

Table 21: Full Model Path Analysis Result

Path Relationships	β	p
DCDifficultyTotal<---Self determination Index	-0.32	0.014

DCDifficultyTotal<---ThoughtProcess	-0.23	0.094
DCDifficultyTotal<--- NewsMediaKnowledge	-0.2	0.114
DCDifficultyTotal<--- SubjectiveNumeracy	-0.11	0.459
DCDifficultyTotal<---Numeracy	-0.12	0.419
DCTotalScore<---StatsCareer	0.26	0.035
DCTotalScore<---DCDifficultyTotal	-0.36	0.007
DCTotalScore<---Self determination Index	0.27	0.043

IV Correlations

SelfdeterminationIndex<- ->ThoughtProcess	0.36
ThoughtProcess<- ->SubjectiveNumeracy	0.30
ThoughtProcess<-->Numeracy	0.24
NewsMediaKnowledge<-->Numeracy	0.22
SubjectiveNumeracy<-->Numeracy	0.48

Squared Multiple Correlations:

Parameter	Estimate	Lower	Upper	<i>p</i>
DCDifficultyTotal	0.33	0.04	0.47	0.05
DCTotalScore	0.34	0.12	0.51	0.011

Multigroup Analysis

The present study extended its analysis to include multigroup comparisons, aiming to uncover potential differences in parameter estimates between gender groups (Hernández-Perlines, 2016). Specifically, the study sought to understand whether the hypothesized relationships within the model held consistently for male and female participants. Significant variations could suggest the

presence of moderation effects, where the association between variables may differ based on gender.

In the initial unconstrained model analysis, where no equality constraints are imposed on the paths across gender groups, the fit indices were suboptimal ($\chi^2/df = 1.681, p < .05$; CFI $< .95$; TLI $< .95$; RMSEA $> .08$), indicating a potential decline in model fit when analyzing gender subgroups separately.

Subsequent analyses introduced constraints on structural weights, covariances, and residuals to the unconstrained model. The non-significant p -values ($p > 0.05$) observed suggest that these constrained models do not significantly deviate in fit from the original, unconstrained model, implying no considerable gender-based differences in path coefficients for variables like disposition, news media knowledge, and numeracy.

Examining the path coefficients more closely, the study noted that subjective numeracy exhibited a non-significant positive effect on perceived difficulty for men ($\beta = .16, p > .05$) and a non-significant negative effect for women ($\beta = -.18, p > .05$). The Self-determination Index (SDI) - indicative of motivation - demonstrated stronger effects on perceived difficulty for men ($\beta = -.60, p < .001$) compared to women ($\beta = -.24, p > .05$). Additionally, the influence of perceived difficulty on comprehension level varied; men showed a significant negative association ($\beta = -.48, p < .05$), whereas for women, the relationship was non-significant ($\beta = -.19, p > .05$). Career aspirations in statistics were also differently weighted, with a more substantial impact on comprehension level for women ($\beta = .36, p < .05$) than for men ($\beta = .15, p > .05$).

In summary, the comparative analysis across gender groups revealed a diminished model fit. Constrained models did not differ significantly from the unconstrained model, as indicated by the invariance of key fit indices (CMIN, CFI, TLI, RMSEA). Nonetheless, when assessing specific path coefficients, the results showed distinct gender-based patterns. Men' confidence in mathematics correlated with higher perceived difficulty, while women' confidence suggested the opposite. The greater impact of motivation on men' perceived difficulty and the stronger influence of career aspirations on women' comprehension level were also notable.

The overall multigroup analysis indicated a relatively poor model fit, potentially attributable to a small sample size. This outcome suggests a direction for future research to focus on larger and more balanced samples across gender groups to confirm these findings and enhance model robustness.

Table 22: Multi-group Analysis Result

Path Relationships	Absolute difference	Male		Female	
	β	β	p	β	p
DCDifficultyTotal<---SDI	-0.358	-0.602	***	-0.244	0.172
DCDifficultyTotal<---ThoughtProcess	0.006	-0.215	0.28	-0.221	0.225
DCDifficultyTotal<---NewsMediaKnowledge	-0.033	-0.166	0.337	-0.133	0.422
DCDifficultyTotal<---SubjectiveNumeracy	0.346	0.163	0.371	-0.183	0.335
DCDifficultyTotal<---Numeracy	-0.141	-0.149	0.419	-0.008	0.965

DCTotalScore<--- StatsCareer	-0.208	0.15	0.399	0.358	0.026
DCTotalScore<--- DCDifficultyTotal	-0.292	-0.478	0.045	-0.186	0.273
DCTotalScore<---SDI	0.022	0.285	0.232	0.263	0.12

IV Correlations

	<i>r</i>
SDI<-->ThoughtProcess	0.304
ThoughtProcess<- ->SubjectiveNumeracy	0.339
ThoughtProcess<- ->Numeracy	0.31
NewsMediaKnowledge<- ->Numeracy	0.184
SubjectiveNumeracy<- ->Numeracy	0.196

Model Fit Indices:

	CMIN	CMIN/DF	GFI	AGFI	CFI	RMSEA	TLI
Unconstrained	50.435	1.681	0.824	0.578	0.639	0.124	0.326
Structural weights	58.305	1.534	0.8	0.622	0.641	0.11	0.471
Structural covariances	62.79	1.281	0.79	0.691	0.756	0.08	0.722
Structural residuals	68.73	1.348	0.769	0.673	0.687	0.089	0.656
Saturated model	0		1		1		
Independence model	112.6	2.011	0.624	0.517	0	0.152	0

Note: ***p<.000

To summarize, the path analysis offers a relatively good model for predicting data comprehension, incorporating results from prior correlation analyses and significant variance tests. The Self-Determination Index (SDI) and career aspirations in statistics positively contribute to data comprehension, whereas perceived difficulty negatively impacts it. A

limitation of the current model is the lack of information on latent variables that could explain perceived difficulty; addressing this could enhance the model's relevance and robustness.

In conducting multigroup analysis across gender groups, the model fit deteriorated, revealing distinct regression relationships between genders. Although the limited sample size may influence these findings, the variations in regression weights are noteworthy. For instance, men's SDI significantly contributes to perceived difficulty, whereas for women, SDI does not show such a relationship. Additionally, men's perceived difficulty significantly affects their data comprehension outcomes, while for women, difficulty does not appear to influence comprehension outcomes. Instead, women's career aspirations in statistics significantly impact their data comprehension, suggesting that men may be more influenced by motivational factors and tend to express their perceived difficulty more directly. In contrast, women are more future-oriented, considering career aspects, and those prepared to use statistics in their careers are better trained in retrieving and comprehending quantitative information.

Chapter 5 Conclusion

5.1 Findings

The study employed a novel approach of combining self-reported measures and eye-tracking devices in order to advance our understanding of how individuals engage with data in news articles.

RQ1: Data Engagement

This research utilized eye-tracking data fixation times to address the first research question which was concerned with how the different data presentations and types impacted audiences' engagement with the data presented in the article. The findings supported that among the three data presentation styles (raw, narrative, and visual), audiences allocated the most fixation time to narrative data and the least to raw data, aligning with the hypothesis that narrative data demands more engagement time than raw data. Notably, visual data consumed less fixation time than expected given the larger area of interest (AOI) for graphs that could possibly require longer engagement periods, as the proportion of participants' fixation on visualization is shorter than the proportion of narrative data. This discrepancy might not directly correlate with comprehension levels; possibly, narrative data's complexity requires prolonged engagement, while visualizations may clarify content, allowing for quicker processing. However, the observed correlation between the time spent on visualizations and self-reported difficulty does not align with current academic perspectives that recognize visuals as enhancing understanding and engagement (Hofer-Pottala, Shannon, 2023). This study's findings suggest that while longer engagement with graphs tends to decrease perceived difficulty, the effectiveness of visualizations is limited if they are not viewed

at all. This indicates that although the principle of communication effectiveness through visuals is valid, it does not fully apply when there is minimal interaction with the data presentations.

Regarding the data types (proportion, statistical elements, relationships, absolute numbers, contextualizers), relational data had the highest fixation times, while absolute numbers attracted the least. This was a reasonable outcome given relational and comparative data's association with larger AOI visualizations like linear graphs. Absolute numbers, represented in smaller AOIs, naturally required less time. Beyond these observations, no additional significant findings emerged from the eye-tracking data analysis. Overall, with the exception of visual data which was lower than expected, the fixation times on the different data presentations and the different data types were fairly consistent with how much of each presentation format and type appeared in the article.

RQ2: Engagement, Comprehension, and Perceived Difficulty

Research question 2 was concerned with the relationship between the different data presentations and types in print news stories with audience comprehension levels (2.1) and perceived difficulty (2.2) and was addressed through eye-tracking data. The study revealed no direct correlation between specific data categories (types or presentations) and overall comprehension, as gauged by whether participants successfully responded to the comprehension questions. However, it was observed that longer engagement with the data in the news article overall correlated with reduced perceived difficulty. Extended time spent on data in visual formats was linked to lower reported difficulty, suggesting that while no specific data category directly influenced perceived difficulty, longer engagement generally facilitated easier comprehension. This aligns with studies

showing that dedicated time spent on a task correlates with improved academic skills, such as math achievements and reading comprehension (Ben-Zeev, Hirsh, Weiss, & Gornstein, 2020).

The findings may indicate more thorough data processing or reflect potential confounds, such as higher interest levels among participants who spent more time reading the data.

RQ3: Participant Dispositions, Comprehension, and Perceived Difficulty

Finally, in research question 3, the association between audience comprehension levels and their dispositional characteristics was explored. The findings indicated that while some demographic factors were associated with some of the dispositional factors (e.g., gender differences in news media knowledge, math grades), they did not directly influence comprehension or perceived difficulty of questions. Two primary dispositional elements emerged as significant predictors of comprehension: an elevated Self-Determination Index (SDI) and an aspiration towards a career in statistics. These factors resonate with Self-Determination Theory (SDT) as proposed by Ryan and Deci (2000; 2023), which emphasizes motivational quality over quantity in promoting outcomes such as psychological well-being, proficient performance, and profound learning. This investigation validates Ryan and Deci's assertion, illustrating that the nature of motivation substantially affects data comprehension.

Career aspirations may serve as an indicator of how individuals engage with data in their daily lives. People who anticipate frequent data interaction tend to prepare for complex data comprehension challenges. This study suggests that such preparation can enhance comprehension levels and reduce perceived difficulty in handling data. Further, Gal's (2002) model on data literacy underscores the importance of a "Critical Stance," alongside "Beliefs and Attitude," as crucial dispositions for acquiring data literacy, asserting that career aspirations and

a motivated mindset significantly contribute to comprehension. Notably, this study elucidates that comprehension is less a matter of demographic and educational background and more about psychological and motivational dimensions.

Additional Findings

An intriguing aspect of the results highlighted differences in how men and women participants perceived and performed on specific comprehension questions, notably questions 4 and 5, which elicited varied responses across gender groups. Question 4 required participants to extract information directly from the article without any need for calculations. In contrast, question 5 necessitated both information retrieval and calculation of a percentage from \$10,000. Women significantly outperformed males on question 4, suggesting a greater aptitude for accurately retrieving information. Conversely, men excelled significantly over women on question 5, indicating a potential proficiency in performing calculations involving percentages. The gender gap in numeracy skills, where females are often outperformed by males, has been noted in several studies (Thomas, Hopwood, & Hatisaru, 2024; Cook, 2018; Suryadarma, 2015). However, recent research suggests that women may excel in graph reading and information retrieval, contradicting earlier findings by Oyeniyi (2013) that did not observe significant gender differences in these areas. Although these results could illuminate gender-specific patterns in graph comprehension, the conclusions are tentative due to the small sample size and uneven gender distribution in the study's participants. These findings call for more comprehensive research to confirm these observations.

In the quest to identify a predictive model, path analysis confirmed that motivation and career aspirations play a pivotal role in comprehension levels. Furthermore, a gender-based comparative analysis was conducted to discern disparities between men and women participants. This analysis uncovered distinct gender-specific patterns: men's confidence in mathematics correlated with a heightened perception of difficulty, whereas for women, increased confidence was associated with a lower perceived difficulty. Specifically, the more confident a man felt about his numeracy skills, the more challenging he found the questions, whereas the more confident a woman was, the less difficulty she perceived. Moreover, the analysis indicated that the more difficult men perceived the questions, the less likely they were to perform well on them. Conversely, the relationship between perceived difficulty and data comprehension was less pronounced for women. Additionally, motivation had a more pronounced effect on men's perception of difficulty, while career aspirations exerted a stronger impact on women's comprehension levels.

These findings contribute an interesting dimension to the current research on the gender gap in adults' numeracy. As Lunardon et al. (2021) suggested, women's numeracy may be more influenced by personality factors than men's. This study further disaggregated "dispositional factors" into motivation (measured by the Self-Determination Index, SDI) and career plans, revealing different levels of significance when considering gender as a moderator. The results indicated that men are more influenced by motivational factors when assessing the difficulty of quantitative information in news articles; the more motivated they are, the less difficulty they perceive. In contrast, women are more career-oriented. Those with clear goals regarding the use

of statistics in their careers, and an understanding of the amount of quantitative information they might encounter, tend to perform better in comprehension than men.

5.2 Discussion

The research illuminates various queries and paves the way for future investigation on the relationship between audiences and media. Eye-tracking technology, as a relatively novel tool for examining reading patterns and attention spans, plays a pivotal role in bridging the gap between gaze behavior and cognitive processing. While the eye-tracking data did not reveal a substantial correlation between engagement and audience comprehension, the study identified some potential patterns in how audience members interact with data in news article and how those strategies related quantitative reasoning and data literacy.

Irrespective of demographic backgrounds, the exploration of dispositional factors revealed several findings. Although participants' knowledge and skills shape their thought processes (Anderson, 1995), it does not reliably predict comprehension of data. Instead, dispositional factors, notably motivation and career aspirations, emerged as crucial determinants of data comprehension in news. This finding challenges the assumption that more knowledge automatically leads to better performance in quantitative analysis and complex problem-solving (Kahneman, 2011). Instead, it highlights the critical role of "willpower" or motivation. The influence of intrinsic and extrinsic motivations on human behavior, including focus and effort, is essential for success in complex tasks (Deci & Ryan, 2000). This underscores the importance of motivation in engaging effectively with challenging problem-solving scenarios. The findings highlight the significance of subjective motivation and life expectations as enhancers of data

literacy engagement in news contexts. This research invites a reevaluation of the presumed neutrality of data and suggests a multidisciplinary approach to understanding data-driven decision-making processes and data literacy in the communication landscape.

The implications of these findings extend to practical applications in the presentation of data and quantitative information within the news media. The results highlight the necessity of crafting content that not only engages and motivates but also clearly guides through problem-solving and has practical, forward-looking applications. An example of a practical approach that utilizes the findings about the impact of motivation and clear guidance in data presentation could be the "Daily Stats" feature in a community newsletter or local news website. This feature could highlight a specific statistic each day, explaining its relevance in everyday life and providing a narrative that ties it back to the readers' daily experiences. For instance, if the statistic of the day is about the percentage of renewable energy usage in the local area, the narrative could include a comparison with the national average, discuss the impact of renewable energy on the environment, and provide tips on how individuals can contribute to increasing this percentage in their daily lives. Such approaches promise to elevate audience engagement by aligning content with audience goals and resonating with their expectations of data in news. Furthermore, these results lay the groundwork for future research aimed at developing a robust model to predict data comprehension in news. This includes refining and expanding upon Gal's (2002) data literacy model to address the intricate theoretical components and quantifying data literacy's application in the news media industry.

5.2 Improvement

5.2.1 Methodological Limitations

The lab setup and research design have some areas for improvement. First, the use of tablets for survey responses led to missed eye movements when participants looked away from the screen, highlighting a limitation in capturing comprehensive engagement levels throughout the duration of the study. Additionally, the absence of a mandatory reading time allowed for significant variability in individual reading durations and engagement with data elements, potentially skewing the data. Future studies might consider integrating screen-based survey questions to maintain within-scope eye movement detection and minimize data variance. While this study adopted the approach of comparing the ratio of data element engagement to total reading time to gauge attention allocation (recognizing that shorter engagement times might reflect specific reading habits rather than diminished attention), controlling for total reading time in future designs could yield more consistent data analysis outcomes.

To further refine the experimental protocol and enhance data quality, a possible adjustment involves a more rigorous calibration process and participant selection criteria. Specifically, only individuals with normal or corrected-to-normal vision should be included in the study. This requirement could be integrated into the participant selection process and explicitly stated in the survey, ensuring the collected eye-tracking data remains accurate and free from deviations. Additionally, participants should be advised against wearing multifocal eyewear during the experiment to avoid complications in eye-tracking calibration, thereby optimizing the accuracy and reliability of the data collected.

Furthermore, Tobii Pro Lab's limitation in handling overlapping Areas of Interest (AOIs) when multiple data presentations or types are tagged necessitated a reliance on ratio analysis to determine engagement with each data category. This method effectively mitigated potential data extraction issues, ensuring a focus on audience attention allocation across different data categories within the study.

In refining the final part of the study design, it is noted that the original approach replaced a conventional post-questionnaire on perceptions of eye-movement recording with Insufficient Effort Responding (IER) questions (Meade & Craig, 2012; Ulitzsch et al., 2024). These questions aimed to gauge participants' attention and effort levels during the experimental session with prompts like, "I put forth 1) almost no, 2) very little, 3) some, 4) quite a lot, 5) my full effort towards the study." To enhance the methodology in the context of eye-tracking technology, it would be advisable to modify these questions to directly address potential biases introduced by participants' awareness of the eye-tracking. Suggested questions by Sülflow et al. (2019) include: "Because of my awareness of eye-tracking, I consciously looked at certain elements" or "While looking at the news article, I forgot about my eye movements being tracked." These questions can help gauge the 'observer effect,' where participants might alter their natural behavior due to the awareness of being monitored (Sülflow et al., 2019). Further, criteria should be established as metrics to identify and potentially discard data that could be skewed by technological intrusiveness or environmental factors. During the data cleaning process, removing such affected data could yield more accurate and representative results, enhancing the reliability and inclusivity of the study's findings.

5.22 Theoretical Limitations

In constructing the path model for predicting data comprehension, there were significant challenges in achieving the right balance between model complexity and simplicity. The goal was to devise a model that not only captures the essence of the collected data accurately, but also avoids the inclusion of superfluous parameters that do not add substantial value. During this process, it became evident that aligning the model with the theoretical frameworks required a more rigorous examination of the links between the independent variables (such as disposition metrics) and the theoretical constructs outlined in the model, for example, the Self-Determination Index's (SDI) association with factors like Personality, Attitudes, or Critical Stance. It should also be noted that we did not have validated tools to measure personality, attitudes, or critical stance with single variable, making it challenging to include these constructs in the model. This shortfall hindered the ability to thoroughly investigate the theoretical foundations of these constructs within the context of the study.

To address these issues, future research should consider integrating direct measurements of these variables into the study design. One potential solution involves designing survey questions as independent variables and indicators for these latent constructs. For instance, to elucidate the relationship between SDI and elements like "Personality", "Critical Stance" and "Belief and Attitude" within Gal's (2002) model, an in-depth review of Self-Determination Theory's (SDT) principles regarding personality development, self-regulation, and nonconscious processes (Deci & Ryan, 2008) is warranted. Developing separate indices to specifically measure metrics related to "Personality", "Critical Stance" and "Belief and Attitude" will allow for a more detailed analysis of their contributions to the model. This approach would not only enhance the

robustness of the SEM but also provide deeper insights into how these theoretical constructs interact and influence data literacy in news consumption. Such advancements would substantially contribute to the field by offering a more nuanced understanding of the dynamics at play of data literacy in news.

An additional theoretical limitation is the absence of a defined standard for what constitutes "perfect reading" of a news article. This absence creates a challenge in setting baseline metrics for evaluating data comprehension and engagement among study participants. Without a clear benchmark, it's difficult to determine the adequate level of fixation on data elements needed for effective perception and comprehension, as highlighted by Glazer (2011). Consequently, it's not feasible to assess whether participants in the study processed the information too quickly or slowly, or whether they sufficiently understood the data elements. This lack of standard complicates drawing definitive conclusions about the participants' engagement and comprehension levels.

Finally, refinements in the research design and adjustments to the laboratory environment, such as expanding the scope of data extraction techniques using Tobii Pro Lab, could significantly enhance the model by providing deeper insights into how often participants visit certain data elements. Incorporating a wider range of eye-tracking metrics—such as pupil diameters, scan paths, heatmaps, and vision clusters, which reveal the areas most frequently revisited—could serve as intermediary predictors of data comprehension (Duchowski, 2017). These additional metrics would offer a more nuanced understanding of participant engagement with and comprehension of data in news articles. Such enhancements would not only improve the

accuracy and depth of the research findings but also contribute to a more sophisticated analysis of how individuals interact with, and process information presented in news media.

References

- Anderson, J. R. (1995). *Cognitive psychology and its implications* (4th ed.). New York: Freeman.
- Arke, E. T., & Primack, B. A. (2009). Quantifying media literacy: development, reliability, and validity of a new measure. *Educational Media International*, 46(1), 53–65.
<https://doi.org/10.1080/09523980902780958>
- Ashley, S., Maksl, A., & Craft, S. (2013). Developing a News Media Literacy Scale. *Journalism & Mass Communication Educator*, 68(1), 7–21. <https://doi.org/10.1177/1077695812469802>
- Attaway, B., Voiklis, J., Barchas-Lichtenstein, J., Hochberg, E., Hammerman, J., Thomas, U. G., LaMarca, N., Santhanam, L., & Parson, P. (2023). Exploring the Relationship between Quantitative Reasoning Skills and News Habits. *Numeracy*, 16(1).
<https://doi.org/https://doi.org/10.5038/1936-4660.16.1.1430>
- Barchas-Lichtenstein, J., Voiklis, J., Attaway, B., Santhanam, L., Parson, P., Thomas, U. G., Isaacs-Thomas, I., Ishwar, S., & Fraser, J. (2022). Number Soup: Case Studies of Quantitatively Dense News. *Journalism Practice*, 0(0), 1–28.
<https://doi.org/10.1080/17512786.2022.2099954>
- Barchas-Lichtenstein, J., Voiklis, J., Santhanam, L., Akpan, N., Ishwar, S., Attaway, E., Parson, P., & Fraser, J. (2021). Better News about Math: A Research Agenda. *Numeracy: Advancing Education in Quantitative Literacy*, 14(1). <https://doi.org/10.5038/1936-4660.14.1.1377>
- Barroso, C., Ganley, C. M., McGraw, A. L., Geer, E. A., Hart, S. A., & Daucourt, M. C. (2021). A meta-analysis of the relation between math anxiety and math achievement. *Psychological Bulletin*, 147(2), 134–168. <https://doi.org/10.1037/bul0000307>

- Beilock, S. L., & Maloney, E. A. (2015). Math anxiety: A factor in math achievement not to be ignored. *Policy Insights from the Behavioral and Brain Sciences*, 2(1), 4-12.
<https://doi.org/10.1177/2372732215601438>
- Ben-Zeev, T., Hirsh, T., Weiss, I., & Gornstein, M. (2020). The effects of high-intensity functional training (HIFT) on spatial learning, visual pattern separation and attention span in adolescents. *Frontiers in Behavioral Neuroscience*, 14, 577390.
<https://doi.org/10.3389/fnbeh.2020.577390>
- Borges-Rey, E. (2016). Unravelling Data Journalism: A study of data journalism practice in British newsrooms. *Journalism Practice*, 10(7), 833–843.
<https://doi.org/10.1080/17512786.2016.1159921>
- Bradshaw, Paul. 2011. *The Online Journalism Handbook: Skills to Survive and Thrive in the Digital Age*. Harlow: Longman
- Bucher, H.-J., & Schumacher, P. (2006). The relevance of attention for selecting news content. An eye-tracking study on attention patterns in the reception of print and online media. *Communications*, 31(3), 347–368. <https://doi.org/10.1515/COMMUN.2006.022>
- Chao, Z. (2023). *Data Journalism: An Area in Need of Deeper Exploration*. In *The Power of Data*. Taylor & Francis Group.
- Cook, R. (2018). Gender differences in adult numeracy skills: what is the role of education? *Educational Research and Evaluation*, 24(1-2), 66-89.
<https://doi.org/10.1080/13803611.2018.1540992>
- Costa Mastrascusa, R., de Oliveira Fenili Antunes, M. L., de Albuquerque, N. S., Virissimo, S. L., Foletto Moura, M., Vieira Marques Motta, B., de Lara Machado, W., Moret-Tatay, C., & Quarti Irigaray, T. (2023). Evaluating the complete (44-item), short (20-item) and ultra-short

- (10-item) versions of the Big Five Inventory (BFI) in the Brazilian population. *Scientific Reports*, 13(1), 7372. <https://doi.org/10.1038/s41598-023-34504-1>
- Cushion, S., Lewis, J., & Callaghan, R. (2017). Data Journalism, Impartiality And Statistical Claims: Towards more independent scrutiny in news reporting. *Journalism Practice*, 11(10), 1198–1215. <https://doi.org/10.1080/17512786.2016.1256789>
- Deci, E. L., & Ryan, R. M. (2000). The “What” and “Why” of Goal Pursuits: Human Needs and the Self-Determination of Behavior: *Psychological Inquiry*. *Psychological Inquiry*, 11(4), 227. https://doi.org/10.1207/S15327965PLI1104_01
- Deci, E. L., & Ryan, R. M. (2008). Self-Determination Theory: A Macrotheory of Human Motivation, Development, and Health. *Canadian Psychology = Psychologie Canadienne*, 49(3), 182–185. <https://doi.org/10.1037/a0012801>
- Duchowski, A. T. (2017). *Eye tracking methodology: Theory and practice*. (3rd ed.). London: Springer.
- Fagerlin, A., Zikmund-Fisher, B. J., Ubel, P. A., Jankovic, A., Derry, H. A., & Smith, D. M. (2007). Measuring numeracy without a math test: development of the Subjective Numeracy Scale. *Medical Decision Making: An International Journal of the Society for Medical Decision Making*, 27(5), 672–680. <https://doi.org/10.1177/0272989X07304449>
- Gal, I. (1994). Assessment of interpretive skills. Summary of working group, Conference on Assessment Issues in Statistics Education. Philadelphia, Pennsylvania.
- 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., & 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(1), 5–28. <https://doi.org/10.1007/s10649-022-10151-7>
- Glazer, N. (2011). Challenges with graph interpretation: A review of the literature. *Studies in science education*, 47(2), 183-210.
- Guay, F., Blais, M. R., Vallerand, R. J., & Pelletier, L. G. (1999). Global Motivation Scale (GMS) [Database record]. APA PsycTests. <https://doi.org/10.1037/t17066-000>
- Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, 31(1), 2–24. <https://doi.org/10.1108/EBR-11-2018-020>
- Hair, J. F., Sarstedt, M., Hopkins, L., & Kuppelwieser, V. G. (2014). Partial least squares structural equation modelling (PLS-SEM): An emerging tool in business research. *European Business Review*, 26(2), 106–121. <https://doi.org/10.1108/EBR-10-2013-0128>
- Hernández-Perlines, F. (2016). Entrepreneurial orientation in hotel industry: Multi-group analysis of quality certification. *Journal of Business Research*, 69(10), 4714–4724. <https://doi.org/10.1016/j.jbusres.2016.04.019>
- Hofer-Pottala, Shannon. (2023). Approaches to Data Visualization in Technical Communication Research: A Systematic Literature Review. Retrieved from the University of Minnesota Digital Conservancy, <https://hdl.handle.net/11299/254252>.
- Huang, J. L., Curran, P. G., Keeney, J., Poposki, E. M., & DeShon, R. P. (2012). Detecting and Deterring Insufficient Effort Responding to Surveys. *Journal of Business and Psychology*, 27(1), 99–114. <https://doi.org/10.1007/s10869-011-9231-8>

- Ikemoto, G. S., & Marsh, J. A. (2007). chapter 5 Cutting Through the “Data-Driven” Mantra: Different Conceptions of Data-Driven Decision Making. *The Yearbook of the National Society for the Study of Education*, 106(1), 105–131. <https://doi.org/10.1111/j.1744-7984.2007.00099.x>
- Jablonka, E., & Bergsten, C. (2021). Numbers don't speak for themselves: strategies of using numbers in public policy discourse. *Educational Studies in Mathematics*, 108(3), 579+. <https://link-gale-com.proxy.bib.uottawa.ca/apps/doc/A685423885/AONE?u=otta77973&sid=bookmark-AONE&xid=57e607d3>
- Jocelyn Raude, Christina Xiao & Pascal Crépey. (2023) Revisiting the primary bias: the role of innumeracy in the misperception of prevalence of chronic illnesses. *Journal of Risk Research* 26:1, pages 64-82.
- John, O. P., & Srivastava, S. (1999). The Big-Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (Vol. 2, pp. 102–138). New York: Guilford Press.
- Joos, M., Rötting, M., and Velichkovsky, B. M. (2003). Die Bewegungen des menschlichen Auges: Fakten, Methoden, innovative Anwendungen [Movements of the human eye: Facts, methods, innovative applications]. In G. Rickheit (Ed.), *Psycholinguistik* (pp. 142-168). Berlin: Stauffenburg.
- Joyce, J. (2021). Bayes' Theorem. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2021). Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/fall2021/entries/bayes-theorem/>

- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87(4), 329–354. <https://doi.org/10.1037/0033-295X.87.4.329>
- Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Straus and Giroux.
- Karaali, Gizem, Edwin H. Villafane Hernandez, and Jeremy A. Taylor. 2016. “What’s in a Name? A Critical Review of Definitions of Quantitative Literacy, Numeracy, and Quantitative Reasoning.” *Numeracy* 9 (1): Article 2. <https://doi.org/10.5038/1936-4660.9.1.2>
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (Fourth edition.). The Guilford Press.
- Knight, Megan. 2015. “Data Journalism in the UK: A Preliminary Analysis of Form and Content.” *Journal of Media Practice* 16 (1): 55–72.
- Koetsenruijter, A. W. M. (2011). Using Numbers in News Increases Story Credibility. *Newspaper Research Journal*, 32(2), 74–82. <https://doi.org/10.1177/073953291103200207>
- Koltay, T. (2015). Data literacy: In search of a name and identity. *Journal of Documentation*, 71(2), 401-415. <https://doi.org/10.1108/JD-10-2013-0128>
- Kovach, B., & Rosenstiel, T. (2014). *The Elements of Journalism: What Newspeople Should Know and the Public Should Expect*. (3rd ed.). New York, NY: Three Rivers Press.
- Krieg, K., Parada-Cabaleiro, E., & Schedl, M. (2022). Do perceived gender biases in retrieval results affect relevance judgements? In *Proceedings of the 2022 Workshop on Algorithmic Bias* (pp. 82-96). Springer. https://doi.org/10.1007/978-3-031-09316-6_10
- Lindsey, L. L. M., & Yun, K. A. (2003). Examining the persuasive effect of statistical messages: A test of mediating relationships. *Communication Studies*, 54(3), 306–321. <https://doi.org/10.1080/10510970309363288>

- Loftus, G. R. (1972). Eye fixations and recognition memory for pictures. *Cognitive Psychology*, 3(4), 525–551. [https://doi.org/10.1016/0010-0285\(72\)90021-7](https://doi.org/10.1016/0010-0285(72)90021-7)
- Lunardon, M., Cerni, T., & Rumiati, R. I. (2022). Numeracy gender gap in STEM higher education: The role of neuroticism and math anxiety. *Frontiers in Psychology*, 13, 856405.
- Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods*, 17(3), 437-455. <https://doi.org/10.1037/a0028085>
- Merriam, A. H. (1990). Words and numbers: Mathematical dimensions of rhetoric. *The Southern Communication Journal*, 55(4), 337–354. <https://doi.org/10.1080/10417949009372802>
- Nadratowski, A., Peters, E., Shoots-Reinhard, B., & Chapman, D. (2022). Objective Numeracy Scale. <https://osf.io/n6a8s/>
- Nguyen, A., & Lugo-Ocando, J. (2016). The state of data and statistics in journalism and journalism education: Issues and debates. *Journalism*, 17(1), 3–17. <https://doi.org/10.1177/1464884915593234>
- Oyeniya, O. (2013). Gender differences in information retrieval skills and use of electronic resources among information professionals in South–western Nigeria. *International Journal of Library and Information Science*, 5(10), 457-465. Retrieved from <https://academicjournals.org/journal/IJLIS/article-full-text-pdf/66836435485>
- Philippe, F. L., & Vallerand, R. J. (2008). Actual environments do affect motivation and psychological adjustment: A test of self-determination theory in a natural setting. *Motivation and Emotion*, 32, 81-89.
- Primack, B. A., Gold, M. A., Switzer, G. E., Hobbs, R., Land, S. R., & Fine, M. J. (2006). Development and Validation of a Smoking Media Literacy Scale for Adolescents. *Archives*

of Pediatrics & Adolescent Medicine, 160(4), 369–374.

<https://doi.org/10.1001/archpedi.160.4.369>

Qualtrics. (2024). About Qualtrics. Retrieved April 10, 2024, from <https://www.Qualtrics.com>

Ramirez, G., Chang, H., Maloney, E. A., Levine, S. C., & Beilock, S. L. (2016). On the relationship between math anxiety and math achievement in early elementary school: The role of problem solving strategies. *Journal of Experimental Child Psychology*, 141, 83–100.

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, 41(1), 203–212. <https://doi.org/10.1016/j.jrp.2006.02.001>

Rivest, L.-P. (1994). Statistical properties of Winsorized means for skewed distributions.

Biometrika, 81(2), 373–383. <https://doi.org/10.1093/biomet/81.2.373>

Robert Brand. (2008) The numbers game: A case study of mathematical literacy at a South African newspaper1, *Communicatio*, 34:2, 210-221, DOI: 10.1080/02500160802456155

Roeh, I., & Feldman, S. (1984). The rhetoric of numbers in front-page Journalism: How numbers contribute to the melodramatic in the popular press. *Text & Talk*, 4(4), 347–368.

<https://doi.org/10.1515/text.1.1984.4.4.347>

Rogers, Simon. (2011). *Facts are Sacred: The Power of Data*. London: Guardian Books.

Rudich-Strassler, A., Hertz-Palmor, N., & Lazarov, A. (2022). Looks interesting: Attention allocation in depression when using a news website – An eye tracking study. *Journal of Affective Disorders*, 304, 113–121. <https://doi.org/10.1016/j.jad.2022.02.058>

Ryan, R. M. (Ed.). (2023). *The Oxford handbook of self-determination theory*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780197600047.001.0001>

- Ryan, R. M., & Deci, E. L. (2000). Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *The American Psychologist*, 55(1), 68–78. <https://doi.org/10.1037/0003-066X.55.1.68>
- Sülflow, M., Schäfer, S., & Winter, S. (2019). Selective attention in the news feed: An eye-tracking study on the perception and selection of political news posts on Facebook. *New Media & Society*, 21(1), 168–190. <https://doi.org/10.1177/1461444818791520>
- Suryadarma, D. (2015). Gender differences in numeracy in Indonesia: evidence from a longitudinal dataset. *Education Economics*, 23(4), 466-486. <https://doi.org/10.1080/09645292.2013.819415>
- Tabachnick, B.G. and Fidell, L.S. (2007), *Using Multivariate Statistics* (5th ed.). New York: Allyn and Bacon.
- Theivendrampillai, S., Cooper, J., Lee, T. et al. Canadian public perceptions and experiences with information during the COVID-19 pandemic: strategies to optimize future risk communications. *BMC Public Health* 23, 796 (2023). <https://doi.org/10.1186/s12889-023-15659-y>
- Thomas, D. P., Hopwood, B., & Hatisaru, V. (2024). Gender differences in reading and numeracy achievement across the school years. *The Australian Educational Researcher*. <https://doi.org/10.1007/s13384-022-00583-8>
- Tobii Pro. (2024). Tobii Pro Spark. <https://www.tobii.com/products/eye-trackers/screen-based/tobii-pro-spark>
- Tobii. (2024). Tobii Academy. Retrieved April 15, 2024, from https://connect.tobii.com/s/tobii-academy?language=en_US

- Ulitzsch, E., Shin, H. J., & Lüdtke, O. (2024). Accounting for careless and insufficient effort responding in large-scale survey data—development, evaluation, and application of a screen-time-based weighting procedure. *Behavior Research Methods*, 56(2), 804–825.
<https://doi.org/10.3758/s13428-022-02053-6>
- University of Ottawa. (2024). ISPR student pool. Retrieved April 15, 2024 from
<https://www.uottawa.ca/research-innovation/inspire/discover-lab/ispr-student-pool>
- Vergara, A., Siles, I., Castro, A. C., & Chaves, A. (2021). The Mechanisms of “Incidental News Consumption”: an Eye Tracking Study of News Interaction on Facebook. *Digital Journalism*, 9(2), 215–234. <https://doi.org/10.1080/21670811.2020.1813047>
- Wang, Z., Lukowski, S. L., Hart, S. A., Lyons, I. M., Thompson, L. A., Kovas, Y., Mazzocco, M. M. M., Plomin, R., & Petrill, S. A. (2015). Is Math Anxiety Always Bad for Math Learning? The Role of Math Motivation. *Psychological Science*, 26(12), 1863–1876.
<https://doi.org/10.1177/0956797615602471>
- Wedel, M., & Pieters, R. (2000). Eye Fixations on Advertisements and Memory for Brands: A Model and Findings. *Marketing Science* (Providence, R.I.), 19(4), 297–312.
<https://doi.org/10.1287/mksc.19.4.297.11794>
- Wheaton, B., Muthen, B., Alwin, D., F., and Summers, G. (1977), "Assessing Reliability and Stability in Panel Models," *Sociological Methodology*, 8 (1), 84-136
- Whitacre, I, Henning, B, Atabaò, ñ: Disentangling the research literature on number sense: Three constructs, one name. vol. 90, issue 1, pp. 95-134. *Review of Educational Research* (2020).
doi: 10.3102/0034654319899706
- Winner, Langdon. 1986. *The Whale and the Reactor: A Search for Limits in An Age of High Technology*. Chicago: University of Chicago Press.

Appendices