CROSS-CULTURAL AUDIENCE COMPREHENSION AND ENGAGEMENT:

A STUDY OF THE EFFECTS OF DATA VISUALIZATION

_____

A Thesis

presented to

the Faculty of the Graduate School

at the University of Missouri-Columbia

_____

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

_____

by

HONG ZHAO

Professor Scott Swafford, Thesis Supervisor

May 2016

The undersigned, appointed by the dean of the Graduate School, have examined the thesis entitled

CROSS-CULTURAL AUDIENCE COMPREHENSION AND ENGAGEMENT:
A STUDY OF THE EFFECTS OF DATA VISUALIZATION

presented by Hong Zhao,

a candidate for the degree of master of arts,

and hereby certify that, in their opinion, it is worthy of acceptance.

Professor Scott Swafford

Professor Sungkyoung Lee

Professor David Herzog

Professor Sean Goggins

ACKNOWLEDGEMENTS

I would like to thank the many international students who shared their time with me and helped me finish the project. Thank you all for your openness.

I especially wanted to thank my graduate committee members for their continuous guidance and kind support along the way. Your wisdom guided me through. This thesis could not have been finished without your great help. I really admire you and love you all deeply.

TABLE OF CONTENTS

Chapter

APPENDIX

LIST OF FIGURES

Figure

Page

LIST OF TABLES

Table

Page

CROSS-CULTURAL AUDIENCE COMPREHENSION AND ENGAGEMENT:
A STUDY OF THE EFFECTS OF DATA VISUALIZATION

Hong Zhao

Professor Scott Swafford, Thesis Supervisor

ABSTRACT

In recent years, data journalism and data visualization has become an emerging and trendy field in online journalism. It is an approach adopted by many news organizations in an attempt to fit for audiences' increasing demand for depth and breadth of information in the "Big Data" era. This research applies the theoretical framework of dual-coding theory and the generative theory of multimedia learning to explore how data visualization influences international audiences' engagement with and comprehension of English-language news stories. The researcher conducted an experiment to compare audiences' comprehension of and engagement with news stories that have data visualizations and those that don't. There were 61 participants, who were international students at the University of Missouri. Each was exposed to two visual and two text stories on climate change from international news organizations The study found that data visualization in news stories enhances audience comprehension and engagement. This study helps justify the frequent and increasing use of data visualization in the news industry and contributes to the broader discussion of multimedia learning in cross-cultural communications.

Chapter 1 Introduction

Data are quickly becoming the driving force in media reformation. An increasing number of news organizations are focusing more attention and resources on digital data projects. Various media organizations have set up their data journalism teams. In 2009, *The Guardian* launched its "Datablog," a section specifically for data reporting and data visualization. Likewise, *The New York Times* has several teams that work on data journalism projects. Its graphics desk with more than 40 staff members is focused specifically on data visualization. Non-traditional media organizations have also joined the movement; the Pew Research Center, ProPublica and Sunlight Foundation all emphasize data analysis in their projects.

Journalists work collaboratively with professionals with knowledge and skills in design, computer science and statistics fields to embrace the "Big Data" era. They dig newsworthy ideas from data, attract audience attention with data visualization projects. Data analysis tools including Google Analytics and Adobe Analytics have also been used to guide media operation and decision-making. "Data have become a torrent flowing into every area of the global economy (p.1)," said Manyika et al. (2011) in McKinsey Global Institute's "Big data: The next frontier for innovation, competition, and productivity" report.

As technological advances facilitate the trends of data journalism, they also contribute to changes in people's reading behavior and the way they consume news. Since the public now plays an active role in the communication – seeking information themselves and even participating in the news-making process – their perspective is an important factor

to consider. The change in the role of the audience has created demands for new forms of media and innovative communication patterns. News stories with data visualizations are a trending type expected to adapt to and satisfy the public. In the information age, people are facing the conflict of infinite information and limited time. Data visualization is based on data, supported by numerical evidence and in the form of illustrations, charts, graphics, maps and interactive elements. News stories with data visualization comb through complicated information, replace lengthy text, use graphics to convey a logical inference, and eventually present a transparent, efficient and accessible news product. This kind of news product helps audiences receive and understand information, encourages them to engage with the stories, and thus optimizes the communication effect.

Increasingly, data journalism as practiced today emphasizes opportunities for audience interaction. For example, *The New York Times* launched a visualization lab that allows readers to create interactive charts, graphs, and maps themselves. Data journalism projects such as ProPublica's "Dollars for Doctors: How Industry Dollars Reach Your Doctors" allowed audiences to search and find information about their personal physicians.

This trend is also taking place worldwide. The project "Where Does My Money Go" by the Open Knowledge Foundation in the United Kingdom drew a huge reaction. It provides data analysis and visualization of information on national government spending. *The Guardian* launched a unique experiment in crowdsourcing, which invited the public to explore and analyze the massive amount of data on the receipts of members of parliament. In China, Xinhua News Agency, the state news agency, as well as almost all leading independent websites, including Sohu, Baidu, Sina, Tencent, and NetEase, have set up their data journalism workshops and have been producing data visualization projects daily.

This research project examined the effects of data visualization on audience comprehension and engagement, thereby to support the decision of news organizations to shift their resources toward data visualization projects. Even though data visualization has gained so much popularity as a news form in recent years, there remains a disproportionally small amount of research on why it has become so popular and how it affects the comprehension and engagement of the audience.

With a more particular approach, this study focuses on the effects of data visualization in international news organizations on global audiences. Non-native speakers of English are the sample group for this study. It is important that international news organizations learn about the effects of data visualization on the comprehension and engagement of global audiences for several reasons. First, global audiences constitute a big proportion of the targeted consumers of international news media organizations. Second, there is a significant amount of research on the impact of visual aids in the field of cross-cultural education and communication. Theoretically, the potential findings will contribute to the broader discussion. Additionally, the study would include a sample of international college students for analysis. College students are a prospective audience of news organizations. They use the Internet as their principal source of news, and the Internet is also the primary platform for data journalism. Students are generally acknowledged to be more likely to accept new things and new ideas, and eager to learn. A better understanding of their likes and dislikes will guide the direction of news outlets. International students form an important part of the current and prospective audiences of news organizations that have global news in mind. While this demographic constitutes a limited sample of the

population of general audiences internationally, the findings should bring insight that could help inform further research on data visualization's impact on more general audiences.

In general, this study concerns the role of data visualization in the comprehension and engagement of audiences who speak English as a second language and consume news from world news organizations including *The New York Times*, *The Guardian*, *The Economist* and *The Washington Post*. Those news media organizations also are leaders in data visualization. The literature review will define the key concepts involved in this study, including data journalism, data visualization, and audience engagement, as well as examine the theoretical framework of the generative theory of multimedia learning and dual-coding theory. The literature review also will provide an overview of the previous studies on the impact of visual aids in cross-cultural communication.

At the practical level, this study will contribute to the industry's understanding of the motivation involving audience engagement and attempt to provide a link between audience engagement preferences and the current popularity of the practices of data journalism. At the theoretical level, to examine the effects of data visualization in news stories, the industry needs academic research that would demonstrate what visualization can accomplish. This study can be used to look at its impact on foreign audiences' comprehension and engagement. It would also be a good reference for future work on examining whether data visualization would change overall audience comprehension and engagement.

Chapter 2 Literature Review

The research on the effects of data visualization on cross-cultural audience engagement extends across several fields, including data journalism, cognitive psychology, and computer graphics and visualization. This literature review attempts to answer some fundamental questions. What is data journalism? What is data visualization? What is an infographic? What is audience engagement? It also includes some studies in the field of journalism that could offer future research directions or topics.

Data Journalism

The concept of data journalism overlaps with those of computer-assisted reporting, investigative journalism, and precision journalism (Meyer, 1979). The industry recognized the first use of computer-assisted reporting in 1952, when the CBS network predicted the outcome of the presidential election with a mainframe computer, even though the results were not put into use. More concrete evidence of use appeared in 1967 when Philip Meyer used survey research to report on a severe city riot for Knight Ridder newspapers. Later in the 1970s, Meyer worked with reporters of the *Philadelphia Inquirer* on sentencing patterns in the local court and with reporters from the *Miami Herald* on property assessment data. Computers became an important tool for news reporting in the late 1980s. In 1989, a story from the *Atlanta Journal-Constitution* on racial disparities in home loans won a Pulitzer Prize. Since then, newsrooms began to involve data in their stories (Houston, 2015; Garrison, 1996, p. 17). The concept of "computer-assisted reporting" these days is more often referred to as "data journalism," as the term "computer" no longer precisely captures the feature of such reporting (Harcup, 2014).

In *A Dictionary of Journalism* (Harcup, 2014), data journalism is defined as:

The mining of databases for trends, discrepancies, and other potential stories. This might involve accessing official information released under freedom of information legislation or provided unofficially by a whistleblower, but it can also include creating a searchable database by inputting into computer programs information gathered by other means. In addition to providing evidence to substantiate (or stand up) a story, data journalism may also enable such stories to be presented graphically and interactively.

A more concise version given by Paul Bradshaw can be found in the *Data Journalism Handbook*: "Data can be the source of data journalism, or it can be the tool with which the story is told — or it can be both (Gray, Bounegru & Chambers, 2012)." Du (2014) defined data journalism as "the use of numerical data in the production and distribution of news." In this study, data journalism is a combination of data reporting and data visualization. Data visualization is a tool that presents stories with data graphically and interactively. It also refers to the products created by such instruments in this study.

Data journalism rose along with the wider phenomenon of big data. Big data is used to describe a massive volume of structured or unstructured data for which the processing exceeds the capacity of traditional databases and software (Beal, n.d.). The amount of data in the world is exploding, and data journalism is one of the solutions in the big data era. Journalists in this role extract meaningful value from big data and present it to the audience. Lewis and Westlund (2014) defined "big data" as "a social, cultural, and technological phenomenon, a complex amalgamation of digital data abundance, emerging analytic techniques, mythology about data-driven insights, and growing critique about the overall

consequences of big-data practices for democracy and society" (p.1). They offered four directions to study journalism in the context of big data: epistemology (journalism's ways of knowing), expertise (journalism's way of doing), economics (journalism's negotiation of value) and ethics (values). Parasie (2014) analyzed the epistemological tensions in the processing of big data by interviewing journalists in the Center for Investigative Reporting on its data-driven project "On Shaky Ground," which took 19 months and an estimated $550,000 to complete. While the process was lengthy and costly, the analysis showed the adjusted knowledge in data can facilitate investigative journalism within a collaborative organization.

There is plenty of research on the use of data in news stories. Studies showed that articles full of data tend to be more credible than articles without data, such as narratives (Koetsenruijter, 2011). Numbers and statistics are more likely to convince the audiences that the conveyed information is trustful (Mayo and Leshner, 2000). The drawbacks include limited audience numeracy, which can reduce reach (Tversky & Kahneman, 1973; Gigerenzer, 2002).

Data Visualization

Visualization is critical to data analysis and data communication. The application of data visualization has always existed in science and technology (Cleveland, 1993). Likewise, in the news field, data visualization has been an emerging presentation format of data reporting. Increasingly, journalists are integrating visualizations into their storytelling.

What is data visualization? Card (2012) defined information visualization as "the use of computer-supported, interactive, visual representations of abstract data in order to

amplify cognition" (p.520). Information visualization combines aspects of information technology, imaging, graphics, scientific visualization, human-computer, as well as human-information interaction (Robertson, Card & Mackinlay, 1993). It is becoming a new trend in the news industry that journalists are now supposed to be equipped with programming knowledge (Wordsman, 2014; Gershon & Page, 2001), statistics and artistic design abilities (Cukier, 2010).

In *A Dictionary of Journalism* (Harcup, 2014), "infographics" is defined as:

A way of presenting information visually online, often involving elements of interactivity whereby website users can click to change which information is displayed, open up videos or other ways of telling the stories contained within the data, or search for particular terms or places.

Although the concept of data visualization is broader than infographics, this study considers infographics as a major example when applied to journalism.

Segel and Heer (2010) categorized narrative visualizations into seven genres based on the number of frames and the ordering of their visual elements: magazine style, annotated chart, partitioned poster, flow chart, comic strip, slide show and video. More often than not, a visualization is comprised of more than one genre. They evaluated successful narrative visualization with the balance between author-driven elements (narrative structure and messaging) and reader-driven elements (interactive exploration and social sharing). As visualization designers and technologists, the researchers focused their analysis on design space and graphical and interactive elements of visualizations, but they gave less attention to the reader-driven experiences.

Most research on data visualization has treated it as independent of the story it accompanies (Kosara & Mackinlay, 2015). To tell a complete story, data visualization often needs to work well with other elements of the story, such as text, video, audio and links. There is a balance to make between data visualization and text in order to tell the story better. While information visualization might provide an intuitive and fast exploration of colossal data resources, it faces the challenge of processing and presenting very complex information. It is also necessary to accommodate different modalities according to the story being told and the audience (Wojtkowski & Wojtkowski).

The effects of data visualization have not been studied much. "Visualization deals with the inhuman scale of the information and the need to present it at the very human scale of what the eye can see," said Martin Wattenberg, a data visualization specialist currently working for Google's "Big Picture" data visualization group (Cukier, 2010). Meaningful metrics to evaluate visualizations would include engagement and interest, ability to remember key points, and information provided to make more informed decisions (Kosara and Mackinlay, 2013).

Cleveland (1994) argued that the construction of a graph involves encoding and decoding processes. Quantitative and categorical information or data are encoded by a display method, and then the visualization is decoded through the human visual system. The choice of information, the graphics technology and the display methods used in encoding should lead to an effective visual decoding. Proper use of display methods can significantly increase the accuracy and efficiency of decoded information. He stated that readers were easily misled when they viewed an enormous amount of information on the

display. The success of a graphical method should be based solely on the amount of information they need to convey.

When a graph is constructed, one essential component to work in close collaboration with is captions. Data visualizations can sometimes be misleading. In this case, captions contribute to a clear understanding of what is graphed. Cleveland (1994) said that a graph and its captions must work together to be understood. Captions should, first, describe everything that is graphed; second, draw attention to the important features of the data; and third, describe the conclusions that are drawn from the data on the graph (p.55). However, in a content analysis by Elzer et al. (2005) of captions in information graphics, the researchers analyzed 100 captions on bar charts and found that captions were often very general and hardly captured the graphic's intended message.

Good data visualization often emphasizes minimizing redundant embellishment to avoid difficulty in reading and focus readers on the conveyed information (Tufte and Graves-Morris, 1983; Wong, 2013). However, Bateman et al. (2010) compared embellished charts with plain ones concerning interpretation accuracy and immediate and long-term recall and found no difference in interpretation accuracy between the two. They also found that visual imagery in graphics makes them more memorable. Even though they only provide a limited number of graphics for analysis, they provide a glimpse of the effects of data visualization projects on human memory and interpretation.

The Effects of Visual Elements in Cross-Cultural Communication

There is little research specifically on the effects of data visualization in the news field while a substantial amount of research focuses on the effects of visual materials on second-language learners' comprehension in the field of cross-cultural education. Therefore, the

study refers to some research on the efficacy of video news. In one of the early studies, MacDonald (1996) studied the effects of using Canadian news videos as English teaching materials in a Japanese university. MacDonald found that visual material gave clues to the information provided by the videos and demonstrated the power of news videos on international issues in shaping social conscience and global consciousness.

Much research has been done on the effects of visual content in facilitating the processing of texts. Fox, Lang, Chung, Lee, Schwartz, and Potter (2004) investigated the comparative amount of comprehension for seven science-related news videotexts in three modified versions that contained computer-generated animations (CGAs), computer-generated texts (CGTs), or no graphics. They found that comprehension was worst for the no-graphics version, although there was little difference between the CGA and CGT versions. Moreover, when the perceived complexity of the news videotext was included as a factor, comprehension was not affected for easier or harder content by the presence of CGAs or CGTs, but more challenging content resulted in significantly less understanding when no graphics were included.

More recently, Cross (2011) explored the role of visual content in Japanese learners' comprehension of news videotexts. Based on dual-coding theory, Paivio (2007) proposed that both visual content and audio content – as in a news videotext – would activate representations in both verbal and nonverbal systems. The researcher combined quantitative and qualitative methods. First, he examined five BBC news videotexts with Walma van der Molen's four-category coding system and quantified their audiovisual characteristics. Ten pairs of participants were exposed to five 90-minutes lessons over five weeks, and then they discussed their understanding of the videotexts. Their dialogue was

audio recorded, transcribed and acted as the unit of qualitative analysis for examining the effects of visual information on their apprehension.

There are some disputes on the effects of multimedia content on audience comprehension. Chandler and Sweller (1991 and 1992) noted the "split-attention" and "redundancy" effects, in which simultaneous multimedia elements might slow mind processing and hinder comprehension because of the extra working memory load. Lang (1995) found that some audiences exposed to multiple elements would direct themselves to only one and intentionally miss the others, given limited short-term memory capacity.

The previous studies of visual content mainly focused on photos, pictures or videos. This study will provide tangible insights into the way data visualizations are processed and how they function in news comprehension.

Image Literacy

Buckingham (2003) said: "Media texts often combine several 'languages' or forms of communication, such as visual images (still or moving), audio (sound, music or speech), and written language" (p. 4). The increasing adoption of data visualization demands higher image literacy as part of readers' overall media literacy. Given the discussion of visual communication effects, readers' understanding of multimedia content relies heavily on individual ability and preferences.

Arizpe et al. (2014) examined how image-based information has an impact on intercultural literacy in childhood education. They found that visual imagery would increase understanding and empathy of children in diverse communities, enhance their enjoyment of reading, and promote analytical and critical thinking. The result of the study suggests that the inclusion of graphic literacy in the classroom is effective and

recommended. Bottrell (2011) also explored the interrelationship between children's image-making (drawing) skill and their cognitive development.

Monnin (2010) analyzed current research on media literacy and justified the use of graphic novels in media literacy education. "Graphic novels" refers to books made up of comic contents. He argued that people are now experiencing a communication revolution, in which readers face a variety of screen-like environments and are required to master both text and image literacy. Graphic novels are not exactly the graphics and data visualization projects discussed in this study, but they have some similarities. Data visualization projects, like graphic novels, also combine texts and images and rely much on image literacy, or rather, graphic literacy. The research on graphic novels or comics is also applicable to visualizations in the news field (Segel & Heer, 2010).

Audience Engagement

Mayer (2011) explored audience engagement from the perspective of journalists and newsrooms and identified three categories: community outreach, conversation and collaboration. Audience engagement involves interaction between journalists and audiences. Journalists try to invite readers to participate in stories through online comments, social media posts and user-generated content. In the digital news era, audience participation has been a popular term in the journalistic study. It involves participation in both the news-making and news dissemination processes. This study will focus on the dissemination phase, in which the two concepts of audience participation and engagement overlap.

Barnes (2014) viewed audience engagement with news content in the digital realm through "the value of engagement to the individual in which pleasure, emotion and a sense

of play are important factors" rather than only "the prism of a contribution's impact on, and value to, the public discourse" (p.130). The rise and popularity of data visualization not only facilitates audiences' comprehension but also satisfies the audience's need for simpler and more entertaining consumption of news.

Xue and Yao (2015) suggested four ways of exploring the application of data journalism: data mining, storytelling, innovative visualization and audience engagement. They argued that there are two ways to enhance audience engagement. One is through interactive design. The interaction happens in the news consuming phase after the stories are complete. Interactive design includes three methods: allowing users to personalize data and data visualization, being open to comment, and providing friendly links for sharing it on social media. The second way to enhance audience engagement is to encourage users to contribute to the news-making process through open sources. News organizations solicit opinions and interests from the public in order to pitch story ideas and encourage the public to participate in the news-making process.

There are a multitude of terms available to describe engagement largely dependent on various contexts. In sociology, Goffman (1974) argued that engagement is close to involvement, which occurs when a person pays attention to and has feeling about a social experience in various contexts. He uses engrossment to describe the degree of immersion attached to involvement in the social experience. In marketing, engagement denotes the degree of attention to an advertising message, and involvement refers to the degree of interest in an advertising message (Zaichkowsky, 1985).

O'Brien and Toms (2010) posited a range of user and system-specific attributes of user engagement in the design of interactive systems including web searching, shopping,

video games and online learning: aesthetics (visual beauty), affect (the user's emotional response), interest (the feeling that accompanies or causes special attention to an object or class of objects), motivation (elements that bring about focus or a desire to proceed with an activity), novelty (features that users find unexpected, surprising, new and unfamiliar), perceived time (users' perception of estimated time spent on the object), focused attention (the concentration of mental activity), challenge (the amount of effort experienced by the participant in performing an online task), control (how in charge users feel over their experience with the technology) and feedback (response or reaction from the task environment or system. Their measures emphasize the user's emotional response and reaction, and the concentration of mental activity.

Lehmann, Lalmas, Yom-Tov and Dupret (2012) analyzed the commonly employed engagement metrics in web usage and provided a model to measure user engagement with online services. They identified three types of engagement metrics: popularity, activity and loyalty. Popularity measures how much a site is used; activity measures how a site is used; and loyalty measures how often users return to a site. The three metrics include time, attention and emotion that users invest. Their model emphasizes the important aspect of interaction and allows initial insights into engagement patterns.

Theoretical Framework

The study will apply dual-coding theory and the generative theory of multimedia learning. This study regards the processing of data visualization such as graphics as an example of the processing of imagery, and applies this theory to model multi-modality journalistic representations. Dual-coding theory was developed by Paivio (1971). It is a psychological theory of cognition and mind that posits that people have two mindsets or

representational systems called verbal and nonverbal systems, processing separately verbal associations and mental imagery. Within the dual-coding theory, the nonverbal system is referred to as an imagery system because its critical functions include the analysis of scenes and the generation of mental images. When people are dealing with environmental information, they build both verbal and visual representations. Later, if people face stimulation evoking their acquired knowledge, they would recall from either memory base or both. This idea is often used to support the importance of imagery in learning and cognitive operations. When both visual and verbal information can be processed, the efficiency in learning process improves (Paivio, 1971, 1990). Dual-coding theory explicitly explains the adaptive functions of the nonverbal and verbal systems, and it sets the fundamental stage for research on imagery. In addition, the theory has inspired much research on the science and practice of educational psychology (Clark, 1991).

Based on the dual-coding theory, Mayer, Steinhoff, Bower and Mars (1995) proposed the generative theory of multimedia learning. The generative theory argues that there are three cognitive steps: First, learners, when exposed to material comprised of both verbal and visual contents select the verbal and visual information independently into their verbal and visual memory bases; second, they organize them into separate verbal and visual representations; and third, they integrate these two modalities into their own presentation (as cited in Plass, Mayer & Leutner, 1998, p. 26). According to the general theory of multimedia learning, learners with more than two kinds of retrieval routes, including verbal and visual representations, can understand better and are more likely to recall the information (Plass, Mayer & Leutner, 1998).

Plass, Mayer and Leutner (1998) summarized two sets of relevant psychological theories of human learning in multimedia learning environments in their study comparing the effects of visual and verbal annotations of German words on the comprehension of English-speaking students. The first were cognitive psychological theories on processing information in visual and verbal modalities. The second represented psychological theories of individual differences and preferences in a multimedia learning environment.

In the study, 103 college students learning German were first classified as verbalizers or visualizers, or neither. They were then asked to read a short story written in German, along with annotations for some marked words in the form of text, pictures and videos. After reading the story on the computer, the students were asked to finish a questionnaire and posttests on vocabulary and comprehension respectively. Each marked word served as a unit of analysis, and the students' choices of annotations (verbal or visual or both) were assigned to each word. The researchers took advantage of a multimedia program that recorded the actions of students, including turning pages and selecting annotation options. In this case, the observation of participants' preferential behavior provides a stronger and more persuasive measure of their individual differences in an authentic situation.

Bateman et al. (2010) compared the effects of embellished data visualization projects with plain ones. They prepared 14 charts, either embellished or plain, and asked 20 college students to read those charts and describe them. Then the 20 students were evenly divided into two groups. The first group was asked to finish an immediate recall task and a preference questionnaire after their description, and the second group completed its recall after two to three weeks. The data Bateman et al. collected included the interpretation of charts, recall results, a user preference questionnaire, and eye gaze. They used paired t-

tests for all the analyses and found that people's interpretation accuracy of embellished charts was no worse than for plain charts, and their long-term recall was significantly better.

The methods in this study are based on those designed by Plass, Mayer and Leutner (1998), Bateman et al. (2010), and Cross (2011), which was discussed earlier in literature review part.

Research Question and Hypothesis

Data visualization is becoming one of the mainstream trends of news media and online journalism. Journalism schools are equipping students with data reporting and visualization techniques. News organizations are hiring people who know programming, statistics and artistic design. Visualizations – including illustrations, charts, maps, and interactive graphics – offer audiences new alternatives to understand stories and encourage them to interact. However, few studies have been done from the perspective of the audience to explain this popularity.

Attracting audiences' attention has always been the goal of journalists. Newspapers and broadcast media want their audiences to spend as much quality time on their products as possible. Online media try to grow their audience and make them stay as long as they can. News organizations such as Reuters, *The New York Times*, *The Washington Post*, *The Guardian* and BBC all set up their Twitter account especially for infographics content, to generate more clicks and shares on social media. Will data visualization applied in news stories help media outlets achieve their objectives? This question is especially relevant, as previous research suggests that the effect of visual elements in cross-cultural communication is significant.

From what has been discussed in the literature review, previous studies foster considerable discussion on the effects of visual elements in the communication and education field. Some research found that imagery facilitates the comprehension process and improves the ability of an audience to consume the news (MacDonald, 1996; Cross, 2011; Fox, Lang, Chung, Lee, Schwartz & Potter, 2004). Image-based methods and graphic novels were found helpful in children's education in intercultural literacy (Monnin, 2010).

This study will take a brief look at the efficacy of data visualization and posits that data visualization facilitates the understanding and engagement of readers from diverse backgrounds. An experiment would be conducted in order to answer the following research questions:

RQ1: Does data visualization in international news stories boost the engagement of diverse audiences who speak English as a second language?

Based on the theoretical framework of dual-coding theory and the generative theory of multimedia learning, visualization is generally more powerful in conveying data-related information, and international audiences would have a better comprehension of the news when presented with both visual elements and texts. That leads to this hypothesis:

H1: Data visualization in news stories will facilitate the comprehension of audiences who speak English as a second language.

Chapter 3 Method

This research was intended to find out whether data visualization in news stories – including illustrations, charts, maps, and other interactive elements – would influence the comprehension and engagement of non-native speakers of English. In order to analyze this topic, this study applied a quantitative research method. This method includes an experiment in which participants who speak English as a second language read four articles (two in full text and two with visual elements) and answered questions that were designed to measure their comprehension of and engagement with English-language news depending on the different formats of news presentations. A comparison was then made between the measured values of comprehension of and engagement with news stories that included or lacked data visualizations.

Participants

Participants were 61 international students who were studying at the University of Missouri, and who were non-native speakers of English. They were recruited through the Intensive English Program and English clubs on campus. According to a priori test, a sample size of 51 participants would be sufficient to achieve a statistical power of 0.8, a probability of 0.05 and anticipated effect size of 0.5 for a within-subject design using a paired t-test. Their IELTS or TOEFL scores and their news-consuming habits were recorded voluntarily as a reference to their language ability. Participants received no prior training or notification on visual focus, nor did they receive any input from the researcher throughout the study, in order to avoid manipulating the direction of their comprehension and engagement. All participants were anonymous. They were not asked to provide their

name or any identifiable information when filling out the survey. The Anonymize Response option of Qualtrics made sure to scrub the response of any identifying personal information, such as the IP address, before saving it in the results.

Table 1			
A summary the participant demographics			
		Count	Percentage
Age	18-20	2	3.3%
	21-24	28	45.9%
	25-30	31	50.8%
Gender	Male	23	37.7%
	Female	38	62.3%
Ethnicity	Hispanic	1	1.6%
	Asian	60	98.4%
Nationality	Chinese	60	98.4%
	Puerto Rican	1	1.6%
Native language	Mandarin	58	95.1%
	Cantonese	2	3.3%
	Spanish	1	1.6%
Years of learning English	1-10	12	19.7%
	11-15	36	59%
	16-20	12	19.7%
	21-25	1	1.6%
Language test scores	TOEFL scores		
	71-80	2	3.3%
	81-90	6	9.8%
	91-100	7	11.5%
	100-110	9	14.8%
	110-120	5	8.2%
	IELTS scores		
	5.5-6	2	3.3%

	6.5-7	4	6.6%
	7.5-8	5	8.2%
	Did not respond	21	34.4%
Education level	Freshman	1	1.6%
	Sophomore	2	3.3%
	Junior	2	3.3%
	Senior	11	18%
	Master's student	41	67.2%
	Doctoral student	4	6.6%
Major	Arts	1	1.6%
	Business	2	3.3%
	Chemistry	1	1.6%
	Communication	2	3.3%
	Economics	2	3.3%
	Education	2	3.3%
	Engineering	21	34.4%
	Finance	2	3.3%
	Journalism	17	27.9%
	Mathematics	1	1.6%
	Psychology	1	1.6%
	Science	2	3.3%
	Statistics	4	6.6%
	Strategic communication	3	4.9%
	Did not respond	1	1.6%

Experimental Design

There were a total of eight stimulus news stories in which half of them had visualization components and half of them didn't. Participants were shown four news stories on climate change: two with data visualizations and two in full text, and answered

22

evaluation questions after each story. The two levels of data visualization were created to represent two story types for each participant. As for the stories with visualizations, the experimental design made sure that each participant was exposed to both interactive and static visualizations.

The presentation of the four stories to participants was random. Within each survey, the Qualtrics program randomized the order, so each saw a different order of the four cells. After reading each story, participants completed a questionnaire measuring their comprehension of and engagement with the story. After completing the exposure to the four stories and answering relevant survey questions, they were asked to answer 11 basic questions about their demographic information and news-consuming habits.

Measures

Measurement of comprehension.

The comprehension variable was measured by asking the participants the open-ended question, "What is the main idea or proposition of this story?" The answers were evaluated and coded by researchers. Participants were asked to explain what the story was about. This was done to examine whether and how data visualization would affect participants' comprehension of the story. In order to ensure reliability of assessing participants' answers, the researcher and a second independent coder evaluated and coded the responses into four categories: Three points (all correct): the participant provides a clear message of what the story is about; Two points (mostly correct): The participant provides most relevant information of the story; One point (mostly incorrect): The participant provides some relevant information but misses most details; Zero points (all incorrect):

The participant provides no correct information; or zero points (I don't know): The participant stated they don't know what the story is about.

Measurement of audience engagement.

Audience engagement was studied as a multidimensional construct; therefore, it was imperative to build a multidimensional survey instrument. The items were derived from attributes identified in the literature review. The metrics of engagement, in this context, considered the simple model of the three metrics commonly used for user engagement with online services: popularity, activity and loyalty (Lehmann, Lalmas, Yom-Tov & Dupret, 2012), and the more detailed attributes of engagement proposed by O'Brien and Toms (2010): affect, interest, motivation, feedback, novelty and perceived time. Not all measures were appropriate for use because the study of O'Brien and Toms was more focused on engagement with a shopping website.

There were nine questions to measure audience engagement: 1) How do you rate the value of the story? 2) How informative do you think the story is? (text stories), or How informative do you think the graphics are? (visual stories) 3) Does this story make you want to read more related stories? 4) Does this story make you want to get more involved in this issue? 5) How would you evaluate the pace at which you read this article? 6) How likely would you be to comment on this story? 7) How likely would you be to recall this information later on? 8) How likely would you be to share this story on social media? 9) How likely would you be to talk about this story to your friends?

Questions 1 – 2 attempted to measure participants' interest in the stories and their perceived importance. Question 3 tried to assess the interest attribute of engagement, seeing if the story would draw particular attention to this subject. Question 4 sought to measure

the motivation attribute of engagement, examining whether reading of the story would bring about focus or a desire to proceed with an activity. Question 5 attempted to evaluate the perceived efficiency of participants and whether the story successfully held their attention. Question 6 aimed to gauge audience engagement with the activity metric and whether the participants would be willing to provide feedback. Question 7 measured engagement with the loyalty metric and the affect attribute of engagement, which involves participants' emotional response to the story. This item also emphasized "continued use," which Jacques (1996) viewed as an outcome of engagement. O'Brien and Toms (2008) proposed the concept of "reengagement," which signifies that engagement has "ebbs and flows" during an individual's interaction with a system (p.53). Questions 8 – 9 evaluated engagement with the activity and popularity metrics. Questions 6, 8 and 9 were also associated with "interactivity," which is more related to online user engagement and necessary to consider in the context of the news stories.

Procedure

The experiment was conducted through Qualtrics, an online survey software. The online survey was active from March 21 to April 3, 2016. It consisted of six blocks (web pages): The introductory page and consent form explained the purpose of the survey, the anonymity, the length of time it would take to complete the survey, and participants' rights to withdraw anytime without consequences from the study. The next four blocks or pages contained four stories and 15 questions relevant to each article. Apart from two open-ended questions, other questions were provided with answers using a five-point Likert scale. Block five contained nine demographic questions. Block six included two questions asking

about participants' news-consuming habits. Respondents went to the next block of the survey by clicking on the arrow button at the bottom right-hand corner of the screen.

All participants completed their sessions at locations convenient to them and without supervision. The entire procedure took from 15 to 45 minutes. A progress bar located at the top of the screen informed participants about their progress through the survey. During the experiment, participants were exposed to two visual stories and two text stories in a random order. As for the two stories with visualizations, each participant viewed one with static graphics and the other with interactive graphics. The inventory for the survey was made up of eight news stories. They were drawn from international news organizations including *The New York Times*, *The Guardian*, *The Economist* and *The Washington Post*. These eight stories were selected based on the following qualifications: theme, length, covered regions and data visualization type. In this study, the eight stories are focused on the same topic, climate change. Climate change is an international issue, which non-native speakers of English would typically consume or be interested in from a source other than their mother language. It is also a topic that involves a mass of data. However, when selecting the news articles, in order to make sure every article remained of the same interest to the participants, the researcher attempted to avoid the influence of immediacy, which is an essential element for news stories. Phrases or sentences that indicated time, such as "Within days the world's leaders will flock to Paris" in Visual D, were deleted from the articles. Also, because the design and operation of the experiment and survey lasted for a while, it was hardly possible to employ the latest news stories for the reading. Fortunately, the new scientific or research findings appeared in the news stories of climate change were relatively less time-sensitive. And the news articles with

data visualizations are more like summaries of the current situation of climate change. Besides, the regions these stories covered were global or had global influence. All the stories except the interactive visualizations were manipulated so as to have a similar length.

Four stories were text only and the other four stories contained mainly visual elements. The "visual" stories were centered on the visual element, which means that if the graphics were taken out, the stories would no longer make sense to readers. The visual stories exposed to participants include both interactive and static data visualization types commonly used in the journalistic field. The interactive data visualization invited participants to explore and interact with the story while the static data visualizations include maps, charts, infographics and illustrations. To make sure the four "text" stories are text only, any photos attached to the stories were deleted in the experiment design. And they were presented directly in the Qualtrics survey. The static graphics were embedded through HTML in the survey. However, due to technical restrictions, for interactive graphics, participants were directed to the websites through links in the survey and finished reading on the original sources.

The data collected from the experiment include a preference questionnaire and experimenter-coded response scores.

Data Preparation

The question that tried to measure the degree of comprehension was an open-ended question, "What is the main idea or proposition of this story?" The entered answers were evaluated by the researcher and a second coder on a scale of zero to three points.

For example, for Text D: "Seas are rising at fastest rate in last 28 centuries," a participant answered, "Sea level is most likely rising faster than any other time in the past

centuries, with the rate of increase growing sharply. Scientists also confirmed that if emissions were to continue at a high rate, the ocean could maintain rapid rise." The participant provided a clear message of what the story is about and his or her answer was coded three points. One participant summarized, "Climate changed significantly since 1990, and most of the time it was on the rise." for Visual B: "Climate change in charts." This response was coded two points because the participant provided most relevant information of the story but missed the global CO_2 emissions part. For Visual C: "How much warmer is your city in 2015," a participant answered, "the world becomes warmer." This was coded one point; the participant provided some relevant information but missed most details. One participant entered, "No idea" for the question on Visual C. This was marked zero points because the participant provided no relevant information.

The second coder received systematic training before the coding process and coded independently scores of participants' responses. If there was any disagreement that appeared in the coding process, the researcher and the coder conducted in-depth discussion until fully understanding the coding rules from the codebook. For any disagreement between the coders on any of the responses, consensus-based annotation was utilized (Ang et al., 2002), in which the coders discussed the answers to try to come to an agreement. The researcher also tested the inter-coder reliability (Scott's pi coefficient = 0.91). It met the basic requirements for content analysis (Craig, 1981). All the nine questions measuring engagement were coded on a 1-to-5-point Likert scale.

This study used SPSS for data analysis. First, Cronbach's alpha, a common reliability coefficient, was used to access the internal consistency of the measures of engagement. The reliability coefficient is 0.924 for the nine items that were meant to

measure the levels of engagement with stories that have data visualizations and those that do not. This was considered a "very good" outcome (DeVellis, 2003), which means the scale was internally consistent, and there was no need to remove any of the questions.

Second, a series of paired t-tests were applied to compare the difference in mean scores between visual and text stories on the variables of audience engagement and comprehension. Third, a series of repeated measures ANOVAs were carried out to examine the effects of language test scores, majors, news-consuming habits, and time spent with the stories, and the interaction effects between the four independent variables and stories types (visual stories and text stories). Lastly, descriptive statistics were provided to analyze the comprehension and engagement of participants from different fields of study.

Chapter 4 Results

Data Analysis for RQ1 and H1

RQ1. Research question 1 asked to compare the level of audience engagement between stories that have data visualizations and those that don't. A paired t-test was calculated to compare the mean value of engagement with visual stories to the mean value of that with text stories. A significant increase in engagement from text to visual stories was found ($t(121) = 3.87$, $p < 0.001$). The mean of engagement with visual stories was 3.40 (sd = 0.92), and the mean of engagement with text stories was 3.00 (sd = 0.93). So data visualization in news stories increased the engagement of international students.

In addition, a paired t-test was calculated to compare the mean value of engagement with interactive visual stories to the mean value of that with static visual stories. No significant difference in audience engagement from interactive to static visual stories was found ($t(60) = 1.31$, $p = 0.20$). The mean of interactive visual stories was 3.48 (sd = 0.96), and the mean of text stories was 3.32 (sd = 0.88).

H1. Hypothesis 1 predicted that data visualization in news stories would facilitate the comprehension of audiences who speak English as a second language. In order to test the hypothesis, a paired t-test was performed for the question that measured the degree of comprehension: "What is the main idea or proposition of this story?" Since the hypothesis predicted a clear direction that visual stories would enhance comprehension, a one-tailed t-test was conducted. A significant increase in comprehension from text stories to visual stories was found ($t(117) = -1.97$, $p = 0.03$). The mean of comprehension of text stories was 1.65 (sd = 1.19), and the mean of

comprehension of visual stories was 1.92 (sd = 1.08). The effect was going in the predicted direction. The comprehension of visual stories was significantly better than that of text stories. This result thus is consistent with the prediction that data visualization in news stories facilitates the comprehension of audiences who speak English as a second language and consume news from world news organizations. Thus, Hypothesis 1 was supported.

Data Analyses Including Moderators

The survey recorded participants' latest language test scores (TOEFL or IELTS), their majors and news-consuming habits, and the time a participant spent with the survey. In order to test if those variables moderated the effect of visualization on engagement and comprehension, further analyses were performed in this section.

In order to test the moderating influence of the four factors: language test scores, major, level of news consumption, and duration of time spent on the survey, on comprehension of and engagement with different stories types, four repeated measures ANOVAs were conducted for engagement and comprehension measurement respectively.

First, participants' recorded TOEFL or IELTS scores were categorized to three groups: low (TOEFL: 0-61, IELTS: 0-5), intermediate (TOEFL: 62-90, IELTS: 5.5-6.5) and high (TOEFL: 91-120, IELTS: 7-9), according to the TOEFL scores interpretation and comparison with IELTS from the ETS official website. Second, the researcher divided the majors into three categories: Science/Engineering, Social Science/Liberal Arts/Education and Business, with reference to the classification of fields of study from a survey by Pew Research Center (2014) on college majors. Third, the news-consuming habit was measured

by the question: How often do you enjoy keeping up with news? (1: Never, 2: Sometimes, 3: About half the time, 4: Most of the time, 5: Always). This measurement was referred to a survey on news attitudes and habits by the Pew Research Center (2012). Lastly, the duration a participant spent with each survey was divided into four groups: the 1st group was below 15 minutes, the 2nd group 15 to 30 minutes, the 3rd group 30 to 45 minutes, and the 4th group above 45 minutes.

A 3 X 2 mixed-design ANOVA was calculated to examine the effects of language level (1: low, 2: intermediate, and 3: high) and story type (1: visual story, and 2: text story) on engagement with stories. No significant main effects or interactions were found. The language level X story type interaction ($F(2,38) = 0.97$, $p = 0.39$), and the main effect for language level ($F(2,38) = 1.80$, $p = 0.18$) were not significant. Audience engagement was not influenced by either language level or its interaction with story type.

A 3 X 2 mixed-design ANOVA was calculated to examine the effects of field of study (1: science/engineering, 2: social science/liberal arts/education, and 3: business) and story type (1: visual stories, and 2: text stories) on engagement with stories. No significant main effects or interactions were found. The field of study X story type interaction ($F(2,57) = 0.09$, $p = 0.92$), and the main effect for field of study ($F(2,57) = 1.00$, $p = 0.38$) were not significant. Audience engagement was not influenced by field of study or its interaction with story type.

A 5 X 2 mixed-design ANOVA was calculated to examine the effects of level of news consumption (1: never, 2: sometimes, 3: about half the time, 4: most of the time, and 5: always) and story type (1: visual stories, and 2: text stories) on engagement with stories. No significant main effects or interactions were found. The news-consuming habit X story

type interaction ($F(4,56) = 0.38$, $p = 0.82$), and the main effect for level of news consumption ($F(4,56) = 0.44$, $p = 0.78$) were not significant. Audience engagement was not influenced by news-consuming habit or its interaction with story type.

A 4 X 2 mixed-design ANOVA was calculated to examine the effects of duration (1: below 15 minutes, 2: 15 to 30 minutes, 3: 30 to 45 minutes, and 4: above 45 minutes) and story types (1: visual stories, and 2: text stories) on engagement with stories. No significant main effects or interactions were found. The duration X story type interaction ($F(3,57) = 1.89$, $p = 0.14$), and the main effect for duration ($F(3,57) = 0.59$, $p = 0.62$) were not significant. Audience engagement was not influenced by duration or its interaction with story type.

In the same way, a 3 X 2 mixed-design ANOVA was calculated to examine the effects of language level (1: low, 2: intermediate, and 3: high) and story type (1: visual story, and 2: text story) on comprehension of stories. A significant story type X language level interaction was present ($F(2,36) = 4.10$, $p = 0.03$). However, the main effect for language level was not significant ($F(2, 36) = 1.68$, $p = 0.20$).
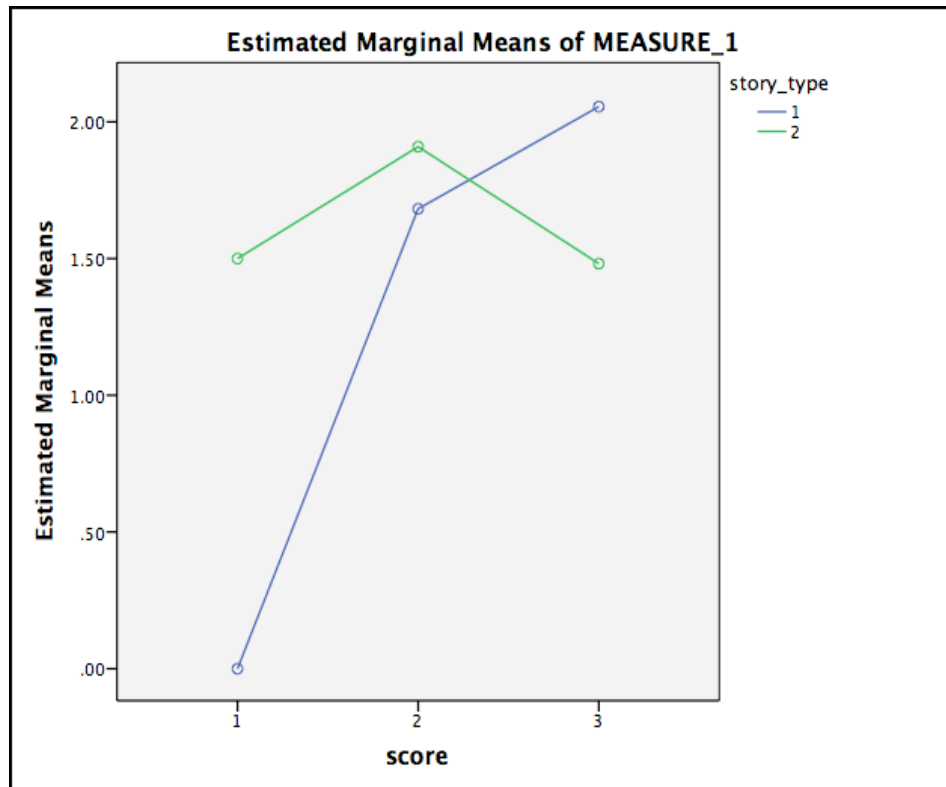
Figure 1: The interaction of language level and story type on comprehension

Table 2		
The means (sd) of comprehension value by three language levels and two story types		
Language level	Story type	Mean (sd)
Low	Visual	0.00 (0.76)
	Text	1.50 (0.74)
Intermediate	Visual	1.68 (0.23)
	Text	1.91 (0.22)
High	Visual	2.06 (0.15)
	Text	1.48 (0.14)

Seen from the profile plot above, participants who have low and intermediate language levels had a better comprehension of stories that do not have data visualizations (story type 2). And for participants who have high language level, their understanding of

visual stories was better than that of text stories.

A 3 X 2 mixed-design ANOVA was calculated to examine the effects of field of study (1: science/engineering, 2: social science/liberal arts/education, and 3: business) and story type (1: visual stories, and 2: text stories) on comprehension of stories. No significant main effects or interactions were found. The field of study X story type interaction $(F(2,54) = 0.31, p = 0.73)$, and the main effect for field of study $(F(2,54) = 0.77, p = 0.47)$ were not significant. Comprehension was not influenced by field of study or its interaction with story type.

A 5 X 2 mixed-design ANOVA was calculated to examine the effects of level of news consumption (1: never, 2: sometimes, 3: about half the time, 4: most of the time, and 5: always) and story type (1: visual stories, and 2: text stories) on comprehension of stories. No significant main effects or interactions were found. The news-consuming habit X story type interaction $(F(4,53) = 0.32, p = 0.86)$, and the main effect for level of news consumption $(F(4,53) = 1.08, p = 0.37)$ were not significant. Comprehension was not influenced by news-consuming habit or its interaction with story type.

A 4 X 2 mixed-design ANOVA was calculated to examine the effects of duration (1: below 15 minutes, 2: 15 to 30 minutes, 3: 30 to 45 minutes, and 4: above 45 minutes) and story types (1: visual stories, and 2: text stories) on comprehension of stories. No significant main effects or interactions were found. The duration X story type interaction $(F(3,54) = 0.40, p = 0.76)$, and the main effect for duration $(F(3,54) = 0.44, p = 0.73)$ were not significant. Comprehension was not influenced by duration or its interaction with story type.

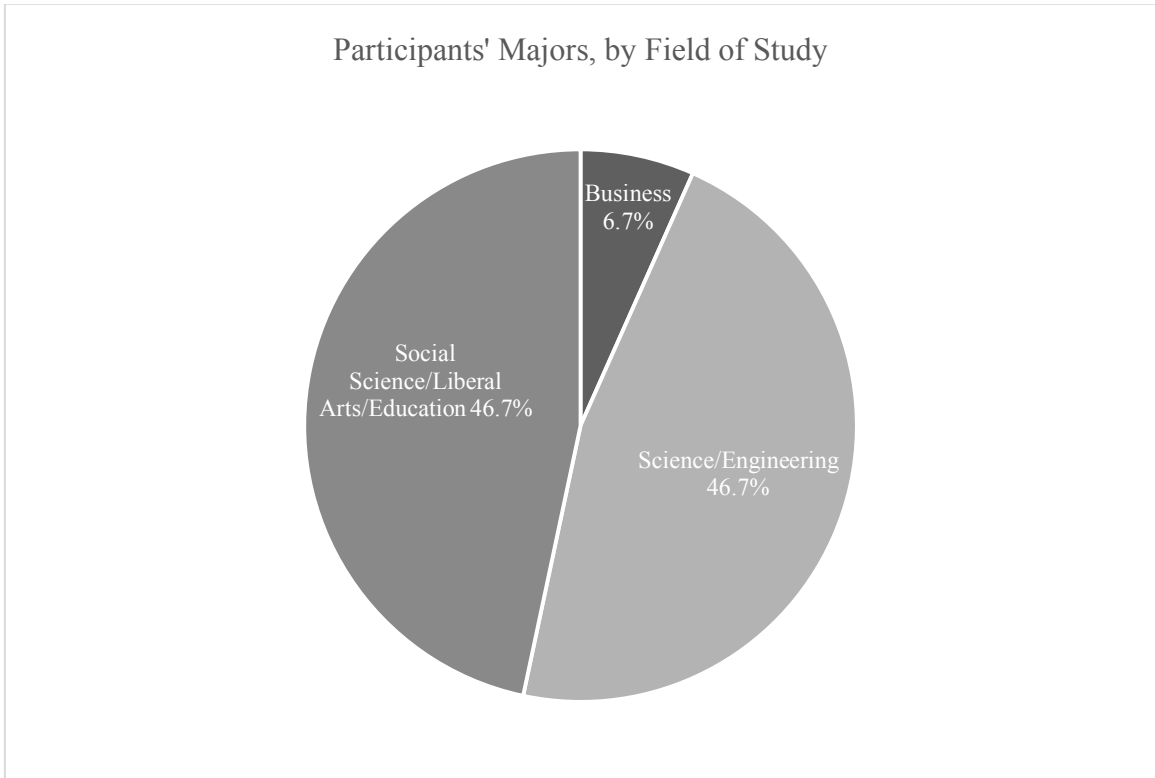**Descriptive statistics analysis of the majors of participants.**

Figure 2: The structure of participants' majors, grouped by field of study

There were four students from the area of business (6.7%), 28 from science/engineering (46.7%), and 28 from social science/liberal arts/education (46.7%).
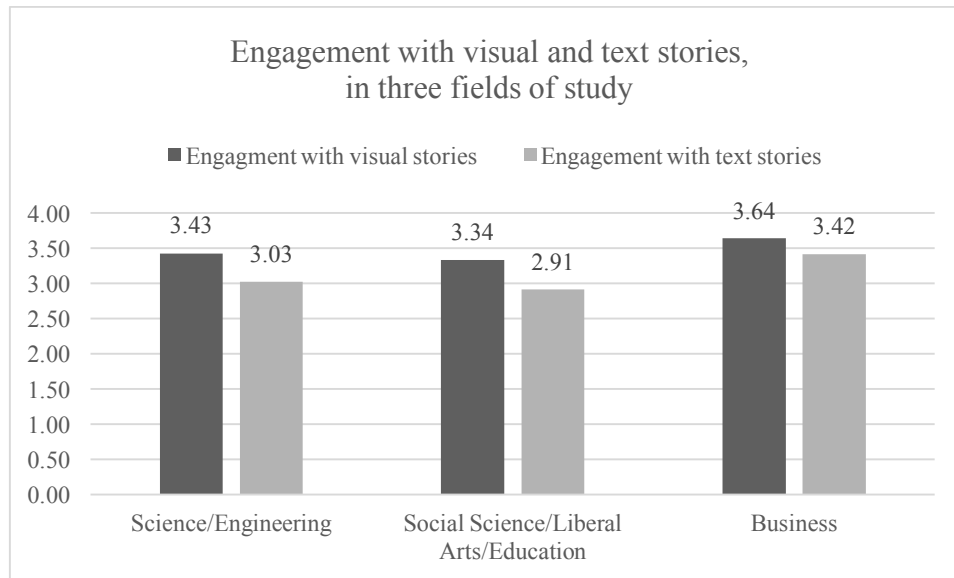
Figure 3: Levels of engagement with stories that contain data visualizations

and those that don't, grouped by field of study.

All students from the three fields of study have a higher degree of engagement with stories that have data visualizations. Students from science/engineering and students from social science/liberal/arts/education majors presented a similar difference between their engagement with visual stories and text stories. The four students from the field of business have the highest degree of engagement with both types of news stories. But the difference of their engagement with visual stories and that with text stories is smaller than that of students from the other two fields.
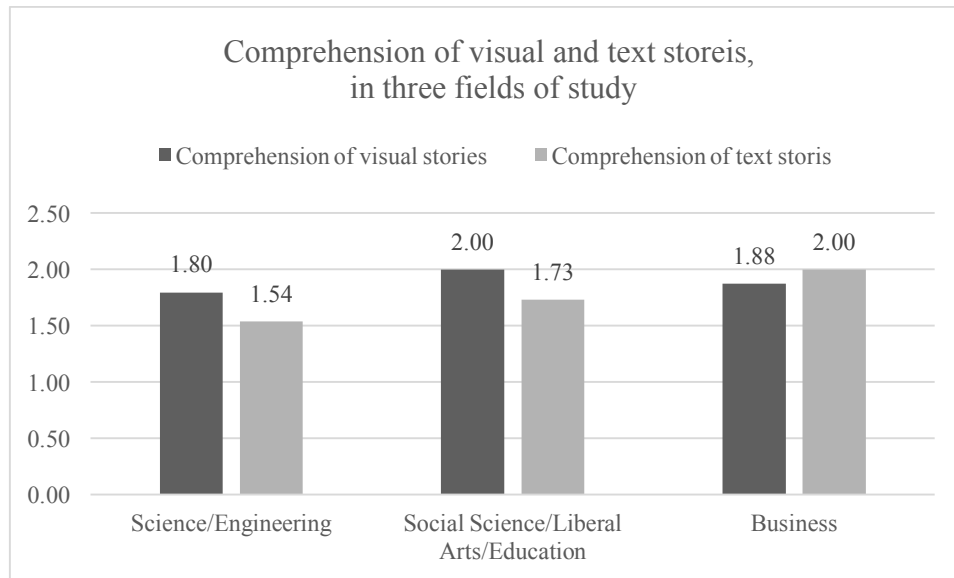
Figure 4: Levels of comprehension of stories that contain data visualizations and those that don't, grouped by field of study.

From the chart above, both students from science/engineering and social science/liberal/arts/education majors have a better understanding of stories that have data visualizations, and they generated a similar difference between their comprehension of visual stories and that of text stories. The four students in the field of business had a higher degree of comprehension with stories that do not have data visualizations.

Despite the fact that these finding cannot be generalized to all students in these fields of study because of the small sample size, this study can make a contribution without moving beyond the particular case studied and was able to provide an initial general understanding of some aspect of the population on the basis of the necessarily limited evidence available in the case.

Chapter 5 Discussion and Conclusion

This study explored the effects of data visualizations on audience engagement and comprehension. The results showed that stories that were focused on visual elements and stories that were in full text led to different levels of engagement and comprehension. This study showed five main findings:

First of all, international college students' engagement with visual stories was significantly better than their engagement with text stories. They tended to attach more value to, show more interest in, and be more willing to interact with news stories that contained data visualizations.

Audience engagement with text and visual stories

■ text ■ visual

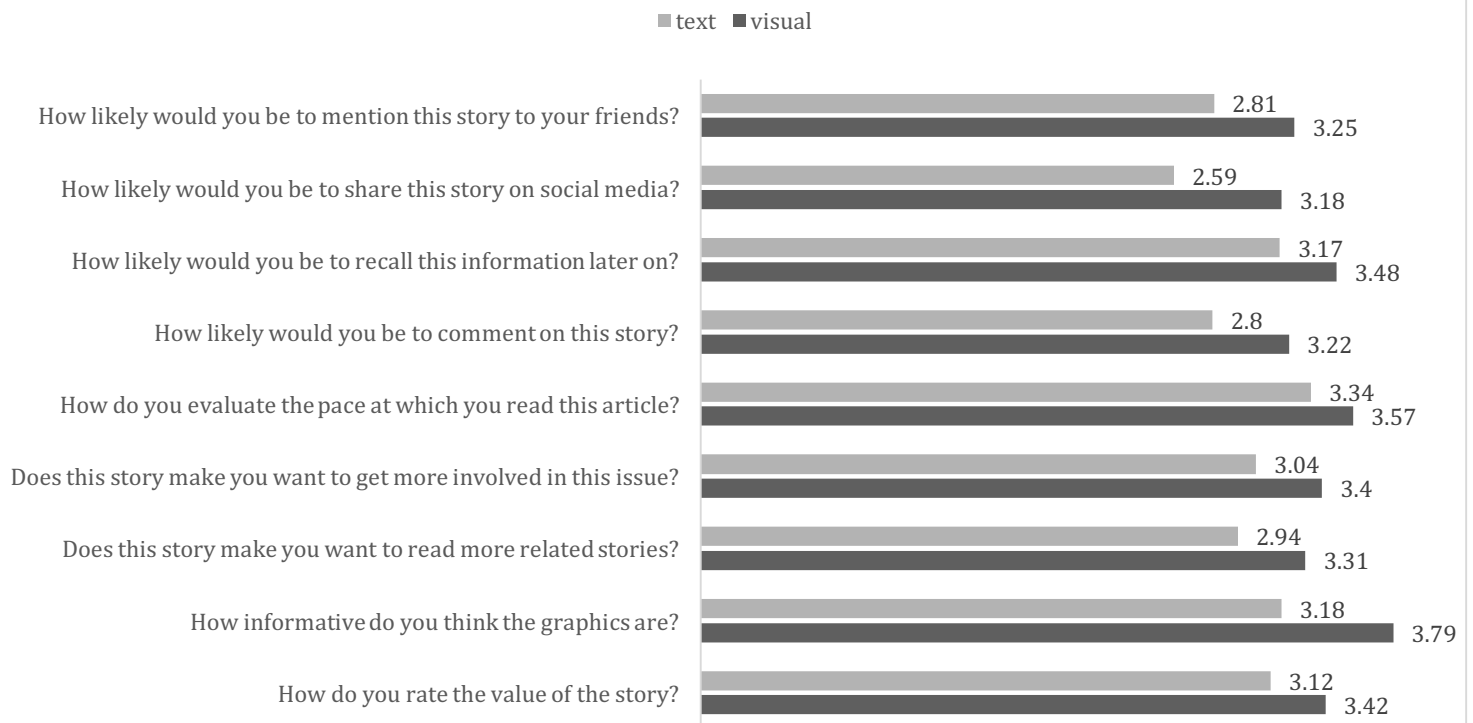| Question	text	visual
How likely would you be to mention this story to your friends?	2.81	3.25
How likely would you be to share this story on social media?	2.59	3.18
How likely would you be to recall this information later on?	3.17	3.48
How likely would you be to comment on this story?	2.8	3.22
How do you evaluate the pace at which you read this article?	3.34	3.57
Does this story make you want to get more involved in this issue?	3.04	3.4
Does this story make you want to read more related stories?	2.94	3.31
How informative do you think the graphics are?	3.18	3.79
How do you rate the value of the story?	3.12	3.42

Figure 5: The questions measuring participants' engagement with
stories that contain data visualizations and those that don't.

Furthermore, international college students' comprehension of visual stories was significantly better than their comprehension of text stories. This finding is in accord with the hypothesis that data visualizations in news stories will facilitate the comprehension of audiences who speak English as a second language.

The language level was found to be a significant factor leading to the difference in comprehension between stories that have data visualizations and those that do not. Students who have low and intermediate language levels had a better comprehension of stories that do not have data visualizations. And students who have high language level has a better understanding of stories that have data visualizations.

Previous studies found that students with higher language proficiency would generate more explanatory inferences than students with lower language proficiency (Zwaan & Brown, 1996). Although the language test scores might not adequately describe participants' current language ability and comprehension skills, they could offer a basic reference in considering language level as a potential factor.

However, majors, news-consuming habits, and the duration they spent on the stories and the survey, had no significant effects on global students' comprehension of the different types of stories. Besides, the three factors, along with language test scores, had no significant effects on their engagement with the two types of stories in this study.

The time participants spent with the survey might depend on their reading pace and their comprehension skills. It was also expected to reflect the engagement of a participant with the story, as "dwell time" on a web page was one important metric to measure engagement (Lehmann, Lalmas, Yom-Tov & Dupret, 2012). Nonetheless, this finding went against the assumption. It's worth nothing that the time participants spent on the survey varied from 13 minutes to more than one day. The participants were able to drop out and then come back to continue to finish it, so this survey allowed the existence of outliers that failed to describe the reading time of the participants. But most participants completed the experiment and survey within an hour and contributed to the valid data.

In addition, there was no significant difference in levels of engagement and comprehension between stories with interactive visualizations and those with static visualizations. Since interactive visualizations provide more opportunities for interactivity, it was predicted that interactive visualizations would lead to greater engagement with the stories. This prediction did not happen in this study. One reason might be that interactive

41

graphics require participants to spend more time exploring the data. They would be unlikely to do this if they were uninterested in the story.

The findings of the study on the effects of data visualization on cross-cultural audience engagement represent a start toward an understanding of how the extensive use of visualizations in news stories influences reader-driven experiences.

As there is a disproportionally small amount of research on the effects of data journalism compared with its popularity and widespread use in the industry, this study was supposed to provide a glimpse of the effects of data visualization on the foreign audiences. In the Big Data era, more research is needed to examine the impact of data journalism. This study contributes to the emerging discussion on data journalism from the perspective of the audience and could help inform news organizations' decision-making.

This study found that international audiences would be more engaged with English-language news stories that include data visualization more than those that do not. These audiences would also understand stories with data visualization better, be more willing to comment on them, share them with friends, and be more likely to recall the information later on.

Despite the favorable impact of data visualizations on foreign audience engagement and comprehension in this study, there is no way to make a clear conclusion that the use of data visualization in news stories will enhance audience engagement and comprehension in general.

It is important to recognize that there are a number of limitations in this research: The participants came from a rather homogenous group because of the researcher's limited access to international students all over the US; the experiment design probably distorted

the natural news-consuming process; there were limited cases of data visualization under discussion.

In addition, Paivio's dual-coding theory also emphasized the importance of individual differences (Paivio & Harshman, 1983). Some research found that multi-modality news presentations might be a burden to audiences due to their limited memory capacity and difficulty of handling more than one form of content at the same time (Chandler & Sweller, 1991&1992; Lang, 1995). Some people choose to focus on one element at a time and block others. Nonverbal processing might be largely dependent on individual preferences and capabilities to process imagery. This reflects somewhat conflicting findings regarding the effects of visual elements.

Besides, the background knowledge of climate change might influence audiences' engagement with and comprehension of the stories. Some studies found that general background knowledge was a significant factor that affects reading comprehensions of second language learners (Haus & Levine, 1985) and familiarity of the topic would predict recall and comprehension of the news story (Tchaicha, 1996).

This study focused on the impact of data visualization on international college students in particular, who represent a large group of international news consumers of international news organizations such as *The New York Times*, *The Washington Post*, *The Economist* and *The Guardian*. Further research and measurement of the effects of data visualization could be done on more general samples of news organizations' targeted audiences. Also, textual analysis of comments can be used to assess audiences' attitudes toward data visualization projects. The most obvious and most important measurement of the effects of data visualization can be the real-time collected data of the web page visits,

time of stay, and the number of comments, and sharing and likes on social media.

REFERENCES

Ang, J., Dhillon, R., Krupski, A., Shriberg, E., & Stolcke, A. (2002). Prosody-based automatic detection of annoyance and frustration in human-computer dialog. Proceedings of the 7th International Conference on Spoken Language Processing (ICSLP). Denver, CO.

Arizpe, E., Bagelman, C., Devlin, A. M., Farrell, M., & McAdam, J. E. (2014). Visualizing Intercultural Literacy: Engaging Critically with Diversity and Migration in the Classroom through an Image-Based Approach. *Language and Intercultural Communication*, *14*(3), 304-321. doi:10.1080/14708477.2014.903056

Bateman, S., Mandryk, R. L., Gutwin, C., Genest, A., McDine, D., & Brooks, C. (2010, April). Useful junk? The effects of visual embellishment on comprehension and memorability of charts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2573-2582). ACM.

Beal, V. (n.d.). Big data. In Webopedia online. Retrieved from http://www.webopedia.com/TERM/B/big_data.html

Gray, J., Bounegru, L., & Chambers, L. (2012). *The data journalism handbook [electronic resource]* (1st ed.). Sebastopol, CA: O'Reilly Media. Available from http://datajournalismhandbook.org/

Barnes, R. (2014). Have You Seen the News? Uncovering the Fan-like Behaviors of the News Audience. *Media International Australia (8/1/07-Current)*, (150), 130-136.

Bottrell, C. (2011). Graphic image-making skills for literacy development. Australian Journal of Language & Literacy, 34(3), 308-321.

Buckingham, D. (2003). *Media education: Literacy, learning and contemporary culture*. Malden, MA: Polity.

Carberry, S., Elzer, S., & Demir, S. (2006). Information graphics: an untapped resource for digital libraries. Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval. New York, NY. doi:10.1145/1148170.1148270

Card, S. (2012). Information visualization. In J. A. Jacko (Ed.), *Human Factors and Ergonomics: Human–Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications* (pp.516-548). Boca Raton, FL: CRC Press.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. Educational Psychology Review, 3(3), 149-210. doi:10.1007/BF01320076

Cleveland, W.S. (1994). *The Elements of Graphing Data*. Summit, NJ: Hobart Press.

Cleveland, W. S. (1993). *Visualizing Data*. Summit, NJ: Hobart Press.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. Cognition and Instruction, 8, 293–332.

Chandler, P., & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology*, 62, 233–246.

Craig, R. T. (1981). Generalization of Scott's Index of Intercoder Agreement. *Public Opinion Quarterly*, 45 (2): 260-264. doi:10.1086/268657

Cross, J. (2011). Comprehending News Videotexts: The Influence of the Visual Content. *Language Learning & Technology*, 15(2), 44-68.

Cukier, K. (2010, February 25). Show Me: New Ways of Visualising Data. *The Economist*. Retrieved from http://www.economist.com/node/15557455

Curtin, P. A., & Maier, S. R. (2001). Numbers in the newsroom: A qualitative examination of a quantitative challenge. *Journalism and Mass Communication, 78(4),* 720-738.

DeVellis, R.F. (2003). Scale development: Theory and application (2nd ed. Vol.26). Thousand Oaks, CA: Sage.

Du, R. (2014). Data Journalism: Mapping the Future. *Journalism & Mass Communication Quarterly, 91*(3), 599-601. doi: 10.1177/1077699014541786b

Elzer, S., Carberry, S., Chester, D., Demir, S., Green, N., Zukerman, I., & Trnka, K. (2005). Exploring and exploiting the limited utility of captions in recognizing intention in information graphics. Proceedings of *the 43rd Annual Meeting on Association for Computational Linguistics*. Stroudsburg, PA. doi:10.3115/1219840.1219868

ETS. (n.d.) Compare Scores. Retrieved from https://www.ets.org/toefl/institutions/scores/compare/

ETS. (n.d.) Interpret Scores. Retrieved from https://www.ets.org/toefl/institutions/scores/interpret/

Fox, J. R., Lang, A., Chung, Y., Lee, S., Schwartz, N., & Potter, D. (2004). Picture this: Effects of graphics on the processing of television news. *Journal of Broadcasting & Electronic Media*, *48(4),* 646-674.

Garrison, B. (2013). *Successful strategies for computer-assisted reporting*. London, UK: Routledge.

Gershon, N., & Page, W. (2001). What storytelling can do for information visualization. *Communications of the ACM*, *44*(8), 31-37.

Gigerenzer, G. (2002). *Calculated Risks: How to Know When Numbers Deceive You*. New York City, NY: Simon & Schuster.

Goffman, E. (1974). *Frame analysis: An essay on the organization of experience*. Cambridge, MA: Harvard University Press.

Gray, J., Bounegru, L., Chambers, L. (2012). *The data journalism Handbook*. Sebastopol, CA: O'Reilly Media. Available from http://datajournalismhandbook.org/

Harcup, T. (2014). Data journalism. In A Dictionary of Journalism. Oxford, UK: Oxford University Press. Retrieved from http://www.oxfordreference.com/view/10.1093/acref/9780199646241.001.0001/acref-9780199646241-e-373

Harcup, T. (2014). Infographics. In A Dictionary of Journalism. Oxford, UK: Oxford University Press. Retrieved from http://www.oxfordreference.com/view/10.1093/acref/9780199646241.001.0001/acref-9780199646241-e-658?rskey=RKk0m7&result=1

Haus, G. J., & Levine, M. G. (1985). The effect of background knowledge on the reading comprehension of second language learners. Foreign Language Annals, 18(5), 391.

Houston, B. (2015, November 12). Fifty years of journalism and Data: A Brief History. Retrieved from http://gijn.org/2015/11/12/fifty-years-of-journalism-and-data-a-brief-history/

Institute of International Education. (2015). Open Doors Report on International Educational Exchange. Retrieved from http://www.iie.org/opendoors

Jacques. R.D. (1996). The nature of engagement and its role in hypermedia evaluation and design. Unpublished doctoral dissertation. South Bank University, London.

Koetsenruijter, A. M. (2011). Using numbers in news increases story credibility. *Newspaper Research Journal, 32*, 74-82.

Kosara, R., & Mackinlay, J. (2013). Storytelling: The Next Step for Visualization. *Computer, 46*(5), 44-50. doi:10.1109/MC.2013.36

Lang, A. (1995). Defining audio/video redundancy from a limited-capacity information processing perspective. *Communication Research, 22,* 86–115.

Lehmann, J., Lalmas, M., Yom-Tov, E., & Dupret, G. (2012) *User Modeling, Adaptation, and Personalization*. Berlin, Germany: Springer Berlin Heidelberg. doi: 10.1007/978-3-642-31454-4_14

Lewis, S. C. & Westlund, O. (2014). Big data and journalism: epistemology, expertise, economics, and ethics. *Digital Journalism*. doi: 10.1080/21670811.2014.976418

MacDonald, E. (1996, February). *Global Awareness through Video News*. Paper presented at the Spring Conference of the Tokyo/West Tokyo Japan Association of Language Teachers, Tokyo, Japan.

Malik, O. (2013, May 10). How the New York Times can fight BuzzFeed & reinvent its future [Web log post]. Retrieved from https://gigaom.com/2013/05/10/how-the-new-york-times-can-fight-buzzfeed-reinvent-its-future/

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A.H. (2011, May). Big data: The next frontier for innovation, competition, and productivity. Retrieved from http://www.mckinsey.com/~/media/mckinsey/dotcom/insights%20and%20pubs/mgi/research/technology%20and%20innovation/big%20data/mgi_big_data_full_report.ashx

Mayer, J. (2011). A Culture of Audience Engagement in the News Industry (Master's thesis). Retrieved from Dissertations and Theses database. (UMI No. 1517401)

Mayo, J., & Leshner, G. (2000). Assessing the credibility of computer-assisted reporting. *Newspaper Research Journal, 21,* 68-82.

McCloud, S. (1993). *Understanding comics: The invisible art*. New York City, NY: Harper Perennial.

Meyer, P. (1979). *Precision journalism: a reporter's introduction to social science methods.* Bloomington, IN: Indiana University Press.

Michael Blastland and Andrew Dilnot. (2010). *The Numbers Game: The commonsense guide to understanding numbers in the news in politics, and in life*. New York, Gotham Books.

Monnin, K. (2010). Teaching Media Literacy with Graphic Novels. *New Horizons in Education*, *58*(3), 78-84.

O'Brien, H. L., & Toms, E. G. (2010). The development and evaluation of a survey to measure user engagement. *Journal of the American Society for Information Science and Technology*, *61*(1), 50-69.

Paivio, A. (1971). *Imagery and verbal processes*. New York, NY: Holt, Rinehart and Winston.

Paivio, A. (1990). *Mental representations: A dual-coding approach.* Oxford, England: Oxford University Press.

Paivio, A. (2014). *Mind and its evolution: A dual coding theoretical approach*. Hove, UK: Psychology Press.

Paivio, A., & Harshman, R. (1983). Factor analysis of a questionnaire on imagery and verbal habits and skills. Canadian Journal of Psychology/Revue canadienne de psychologie, 37(4), 461.

Parasie, S. (2014). Data-driven revelation? Epistemological tensions in investigative journalism in the age of "big data". *Digital Journalism*. doi: 10.1080/21670811.2014.976408

Pew Research Center. (2012). Section 3: News Attitudes and Habits. In Changing News Landscape, Even Television is Vulnerable: Trends in News Consumption: 1991-2012. Retrieved from http://www.people-press.org/2012/09/27/section-3-news-attitudes-and-habits-2/

Pew Research Center. (2014). Usefulness of Major, by Field of Study. The Rising Cost of Not Going to College. Retrieved from http://www.pewsocialtrends.org/2014/02/11/the-rising-cost-of-not-going-to-college/sdt-higher-education-02-11-2014-0-08/

Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (1998). Supporting visual and verbal learning preferences in a second-language multimedia learning environment. *Journal of Educational Psychology*, *90*(1), 25-36. doi:10.1037/0022-0663.90.1.25

Robertson, G. Card, S. and Mackinlay, J. (1993). Information Visualization Using 3D Interactive Animation. Communications of the ACM vol. 36(4), 57-71

Segel, E., & Heer, J. (2010). Narrative visualization: telling stories with data. *IEEE Transactions On Visualization and Computer Graphics*, *16*(6), 1139-1148. doi:10.1109/TVCG.2010.179

Scott, J., & Craig-Lees, M. (2010). Audience Engagement and its Effects on Product Placement Recognition. *Journal of Promotion Management*, *16*(1/2), 39-58. doi:10.1080/10496490903571803

Scott R. M. (2002). Numbers in the News: A Mathematics Audit of a Daily Newspaper. *Journalism Studies, 3(4),* 507-519.

Scott, R. M. (2003). Newsroom Numeracy: A Case Study of Mathematical Competence and Confidence. *Journalism and Mass Communication, 80(4)*, 921- 936.

Tchaicha, J. D. (1996). Background knowledge and second language proficiency: How adult ESL learners make sense of United States news programming. *Dissertation Abstracts International Section A, 57,* 1022.

Tufte, E. R., & Graves-Morris, P. R. (1983). *The visual display of quantitative information, 2(9).* Cheshire, CT: Graphics press.

Tversky, A. & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin, 76(2),* 105-110.

Wilkinson, L. (1999). *The Grammar of Graphics*. New York, NY: Springer Publishing.

Wojtkowski, W., & Wojtkowski, W. G. (2002, October). Storytelling: Its role in information visualization. Proceedings from *The Fifth European Systems Science Congress*. Crete, Greece.

Wong, D. M. (2013). *The Wall Street Journal Guide to Information Graphics: The Dos and Don'ts of Presenting Data, Facts, and Figures*. New York City, NY: W.W. Norton & Company.

Xue, G., & Yao, Y. (2015). 数据新闻的讲故事之道[The storytelling of data journalism]. 新闻与写作[*News and Writing*] (5). Retrieved from http://www.vccoo.com/v/b1095b

Zwaan, R. A., & Brown, C. M. (1996). The influence of language proficiency and comprehension skill on situation-model construction. *Discourse Processes, 21(3)*, 289-327. doi:10.1080/01638539609544960

Table 3: An overview of the titles, publishing dates, types of data visualizations, and media sources of the news stories studied.

News item	Title	Publishing Date	Media organization	Data visualization type	Link
Visual A	Gauging a warming world	Nov. 30, 2015	The Washington Post	Illustrations, infographics	https://www.washingtonpost.com/graphics/national/gauging-climate-change/
Visual B	Climate change in charts: from record global temperatures to science denial	Feb. 1, 2016	The Guardian	Charts	http://www.theguardian.com/environment/planet-oz/2016/feb/02/climate-change-in-charts-from-record-global-temperatures-to-science-denial
Visual C	How Much Warmer Was Your City in 2015?	Feb. 19, 2016	The New York Times	Interactive graphic	http://www.nytimes.com/interactive/2016/02/19/us/2015-year-in-weather-temperature-precipitation.html#ashland_mo

Visual D	Climate change: The state of the planet	Nov. 25, 2015	The Economist	Charts, maps and infographics	http://www.economist.com/blogs/graphicdetail/2015/11/graphics-climate-change
Text A	Climate scientists worry about the costs of sea level rise	March 2, 2016	The Guardian	N/A	http://www.theguardian.com/environment/climate-consensus-97-per-cent/2016/mar/02/climate-scientists-worry-about-the-costs-of-sea-level-rise
Text B	What the Earth will be like in 10,000 years, according to scientists	Feb. 8, 2016	The Washington Post	N/A	https://www.washingtonpost.com/news/energy-environment/wp/2016/02/08/what-the-earth-will-be-like-in-10000-years-according-to-scientists/
Text C	Forests and climate change: Hope for the trees	Nov. 28, 2015	The Economist	N/A	http://www.economist.com/news/international/21684140-modest-progress-has-been-made-saving-forestsit-needs-accelerate-hope-trees

Text D	Seas Are Rising at Fastest Rate in Last 28 Centuries	Feb. 22, 2016	The New York Times	N/A	http://www.nytimes.com/2016/02/23/science/sea-level-rise-global-warming-climate-change.html?ref=topics

News Articles (Note: The articles might be manipulated to have similar lengths.)

Text A

Climate scientists worry about the costs of sea level rise

By John Abraham | The Guardian

As humans add greenhouse gases to the atmosphere, it not only warms the planet, but also raises the oceans. Ocean waters are rising for a number of reasons including thermal expansion of water (as water warms, it expands to a larger volume), as well as ice melt which then flows as liquid into the ocean.

A paper was just published by Drs. Boettle, Rybski and Kropp that dealt with this question. The authors of this study note that if you are concerned about societal and economic costs, the rate of sea rise isn't the entire story. Much of the damage is caused by extreme events that are superimposed on a rising ocean. Damage is highly nonlinear with sea rise.

To explain this, let's think about flooding. Consider a river that has a dike system capable of confining a rise of water up to six feet. Such a system would have little or no economic/societal damage for "floods" up to six feet, but just one more foot of water rise would put the waters over the dike and could cause significant losses. So what really matters is, do events overshoot some level that commences damage?

How does this relate to climate change? Well as we warm the planet we are raising the baseline level of water from which extremes happen. Second, we are making some extreme weather events more likely. To measure the changes to extreme events in the future, the authors use a statistical method to estimate economic losses from coastal flooding. Using Copenhagen and other locations as test cases, they found that economic losses double when water rises only 11 cm. They also find that the costs rise faster than sea level rise itself. So, if we expect a linear increase in sea level over the next century, we should anticipate costs that increase more rapidly.

The authors also look at what are called "tail events" of storm surges. These are unusual events that can cause a large fraction of losses. Superstorm Sandy is an example; the storm surge from that event was very extreme and cause more loss than the combination of many smaller storm surge events.

Text B

What the Earth will be like in 10,000 years, according to scientists

By Chris Mooney | The Washington Post

A large group of climate scientists has made a bracing statement in the journal Nature Climate Change, arguing that we are mistaken if we think global warming is only a matter of the next 100 years or so — in fact, they say, we are locking in changes that will play out over as many as 10,000 years.

"The next few decades offer a brief window of opportunity to minimize large-scale and potentially catastrophic climate change that will extend longer than the entire

history of human civilization thus far," write the 22 climate researchers, led by Peter Clark, from Oregon State University.

The researchers' key contention is that we have been thinking about climate change far too narrowly by only projecting outward to the year 2100, which the research says "was originally driven by past computational capabilities." Rather, we should consider that the long-term consequences of human emissions for global temperatures and sea level will play out over many millennia.

"It's a statement of worry," said Raymond Pierrehumbert, a geoscientist at Oxford University and one of the study's authors. "And actually, most of us who have worked both on paleoclimate and the future have been terrified by the idea of doubling or quadrupling CO2 right from the get-go."

"In hundreds of years from now, people will look back and say, 'Yeah, the sea level is rising; it will continue to rise; we live with a constant rise of sea level because of these people 200 years ago that used coal, and oil and gas,'" said Anders Levermann, a sea-level-rise expert at the Potsdam Institute for Climate Impact Research and one of the paper's authors. "If you just look at this, it's stunning that we can make such a long-lasting impact that has the same magnitude as the ice ages."

The key reason for this is that carbon dioxide stays in the atmosphere for a very long time before being slowly removed again by natural processes. "A considerable fraction of the carbon emitted to date and in the next 100 years will remain in the atmosphere for tens to hundreds of thousands of years," the study noted. Meanwhile, the planet's sea levels adjust gradually to its rising temperature over thousands of years.

So what will the world look like in 10,000 years, thanks to us? That really depends on what we do in the next few hundred years with the fossil fuels to which we have relatively easy access. It also depends on whether or not we develop technologies that are capable of pulling carbon dioxide out of the air on a massive scale, comparable to the amount that we're currently emitting.

Text C

Forests and climate change: Hope for the trees

The Economist

Until the 1960s, forest-clearing accounted for most anthropogenic carbon emissions. Now it causes around 10%—a decline that led many at the UN climate summit in Paris to focus their efforts elsewhere. Though Norway, Germany and Britain said they would make a billion dollars a year available for averting tropical deforestation until 2020, America, France and Japan refused to chip in. Australia trumpeted a pro-tree plan of its own, but has not pledged more money for it. There was little mention of Indonesia's devastating wildfires, or of a 16% uptick in deforestation in Brazil.

Yet the 10% share hugely understates the importance of forests to the fight against climate change. Just as shrinking forests contribute to global warming, growing ones can counter it. During the 2000s tropical forests are estimated to have sopped up and stored carbon equivalent to 22-26% of carbon-dioxide emissions from human activity. Ending tropical deforestation and letting damaged forests recover could cut net emissions by almost a third, creating a space for industrial emissions to fall more slowly.

The Paris agreement failed to create mechanisms, such as carbon markets, that could generate the much larger sums necessary for conservation on that scale. That was expected; on a more realistic measure of progress, forests did pretty well. Reducing deforestation and forest degradation—REDD+, in the jargon—has finally been enshrined as a mainstream climate policy. Over 60 countries included it in their commitments. It got its own clause in the final deal, with an approving nod to ways of designing and running REDD+ schemes agreed on in previous climate talks. Schemes can be funded publicly or privately, with payment for success in leaving trees standing.

Whether REDD+ can fulfil its potential remains unclear. By one measure, the net global deforestation rate has fallen in recent years. Yet, despite improvements in monitoring tropical deforestation and establishing baselines against which it can be measured, REDD+ probably played a minor part. Brazil dislikes the idea of "offsetting"—letting other countries or even firms emit more in return for paying to keep its trees standing. Indonesia has, in effect, downgraded the REDD+ agency it set up in a deal with Norway, and failed to spend most of the billion dollars that the Norwegians provided.

Rich countries say this shows that what is mainly lacking is political will. Those with tropical forests retort that the failure to cough up the large sums they were once promised vindicates their wariness. Both are right: for now, REDD+ is almost as notional as it is necessary.

Text D

Seas Are Rising at Fastest Rate in Last 28 Centuries

By Justin Gillis | The New York Times

The worsening of tidal flooding in American coastal communities is largely a consequence of greenhouse gases from human activity, and the problem will grow far worse in coming decades, scientists reported.

Those emissions, primarily from the burning of fossil fuels, are causing the ocean to rise at the fastest rate since at least the founding of ancient Rome, the scientists said. They added that in the absence of human emissions, the ocean surface would be rising less rapidly and might even be falling.

The increasingly routine tidal flooding is making life miserable in places like Miami Beach; Charleston, S.C.; and Norfolk, Va., even on sunny days.

Though these types of floods often produce only a foot or two of standing saltwater, they are straining life in many towns by killing lawns and trees, blocking neighborhood streets and clogging storm drains, polluting supplies of freshwater and sometimes stranding entire island communities for hours by overtopping the roads that tie them to the mainland.

Such events are just an early harbinger of the coming damage, the new research suggests.

"I think we need a new way to think about most coastal flooding," said Benjamin H. Strauss, the primary author of one of two related studies released. "It's not the tide. It's not the wind. It's us. That's true for most of the coastal floods we now experience."

In the second study, scientists reconstructed the level of the sea over time and confirmed that it is most likely rising faster than at any point in 28 centuries, with the rate

of increase growing sharply over the past century — largely, they found, because of the warming that scientists have said is almost certainly caused by human emissions.

They also confirmed previous forecasts that if emissions were to continue at a high rate over the next few decades, the ocean could rise as much as three or four feet by 2100.

Experts say the situation would then grow far worse in the 22nd century and beyond, likely requiring the abandonment of many coastal cities.

Visual A

Gauging a warming world

By Bonnie Berkowitz, Patterson Clark, Todd Lindeman, Denise Lu, Kevin Uhrmacher and Lazaro Gamio | The Washington Post

https://www.washingtonpost.com/graphics/national/gauging-climate-change/

Visual B

This message contains graphics. Please click here to view from the original source.

Climate change in charts

By Graham Readfearn | The Guardian

Much has been written about climate change in recent months, what with that record-breaking hot year we just had and the qualified success of the Paris climate talks.

But if there's one criticism I'd have of the media coverage, it's this.

Not enough graphs.

So here are four that you might have missed, but that tell us a few things about the state of the climate.

Chart showing average global temperatures from 1850 to 2015 according to three major datasets Photograph: Met Office, UK

So how have global temperatures been recently? You know, while human-caused global warming has been on that pause that didn't happen.

Above is a chart from the UK's MetOffice showing the three main global temperature data sets. These, for the hardcore among you, are known as HadCrut4 (MetOffice), GISTEMP (NASA) and MLOST (NOAA).

Pretty clear, right? All three show 2015 was the hottest year on record.

The chart shows "anomalies" (that's the temperatures above or below a long-term average), in this case, against the average temperature between 1961 and 1990.

So for the MetOffice data, 2015 was 0.75C above the long-term average.

In December 2014, the MetOffice guessed (and when I say guess, I mean they used some powerful, skilful and sophisticated modelling) that 2015 would be between 0.52 and 0.76C above average. So they were right. Just.

Chart showing the MetOffice estimate of global temperatures in the upcoming five years to 2020 Photograph: Met Office

And while we're talking about forecasts from the MetOffice, here's another one of theirs from a few days ago.

The MetOffice now makes "decadal" forecasts – that is, another of their guesses (same caveats apply as before with the term "guess") on "near term" conditions globally over the next five years.

The black line shows actual temperature measurements. The red band are previous predictions and the green band shows the expected range according to climate models.

In short, according to the MetOffice, the forecast is "for continued global warming largely driven by continued high levels of greenhouse gases".

Chart showing the rising heat content of the world's oceans Photograph: NOAA

You can't really talk about global warming without talking about the oceans, given that this is where the vast majority of the extra heat and the extra carbon dioxide ends up.

The above chart, from the National Oceanic and Atmospheric Administration, shows the latest temperature readings from ships and from a network of almost 4,000 floats – known as the Argo array.

A study published a couple of weeks ago found heating of the oceans was accelerating, particularly in the deeper ocean.

Chart showing global CO2 emissions from fossil-fuel use and industry since 1990 and emissions intensity CO2 /GDP. The red symbols are projections for 2015. Photograph: Nature Climate Change

Rising CO2 emissions and economic growth have been like two kids who met as childhood sweethearts and just never broke up. It has always seemed like the two were made for each other.

But if you want to keep an economy growing but tackle the critical issue of rising emissions, then there has to be a break-up. Decoupling economic growth from rising emissions is the big challenge.

This chart comes from a December 2015 paper, also in the journal Nature Climate Change.

The black squares show the amount of greenhouse gases being emitted for every unit of GDP. Clearly, that's been dropping. The white circles are carbon dioxide emissions, and clearly they've been going up.

Success would mean both those lines heading rapidly south. The study showed that for the first time, emissions had dropped slightly while the world's economy grew.

So is there a break-up on the cards?

Commenting on the finding, Myles Allen, professor of geosystem science at the University of Oxford said: "Is this the beginning of the end of global warming? Probably not. But let's hope it is the end of the beginning."

Visual C

How Much Warmer Was Your City in 2015?

By K.K. Rebecca Lai | The New York Times

http://www.nytimes.com/interactive/2016/02/19/us/2015-year-in-weather-temperature-precipitation.html#ashland_mo

Visual D

This message contains graphics. Please click here to view from the original source.

Climate change: The state of the planet

BY M.S.L.J. and the data team | The Economist

None would deny the consequences of the greenhouse effect on the Earth's temperature, the role carbon dioxide plays therein or humanity's part in adding to the level of that gas in the atmosphere over the past few centuries. The charts below help represent the state of climate change.

The world is already 0.75°C warmer than it was before the Industrial Revolution. Two recent studies, one published in Science and another in Climatic Change, suggest that a much-debated hiatus in global warming between 1998 and 2012 in fact never happened. The first says the cooler readings were caused by a switch to measuring ocean temperatures from buoys rather than ships; the second finds that the statistical tools used to demonstrate the apparent slowdown were inadequate.

Meanwhile, the atmospheric concentration of carbon dioxide has risen from just under 340 parts per million (ppm) in 1980 to 400ppm today. Many are seeing the effects of climate change in the fact that 2015 has been the hottest year around the world since records began—the strong Niño, a Pacific-wide climatic phenomenon, was helping too.

The Arctic is heating up twice as quickly as the rest of the planet, and is the region where the impact of a warming climate can most vividly be seen. While a 1°C rise in temperatures at the equator will have some noticeable effects, in the Arctic such an increase melts the ice. Summer sea-ice cover has declined more than 40% over the past 36 years, causing huge changes both to local biology and to global meteorology. Since America bought Alaska from Russia in the 19th century, the average sea level there has

risen by more than 20 cm and coastal villages in the state are threatened by the lapping of chilly waters. The sea is expected to rise much further by 2100. Warm periods in the past 3m years almost certainly saw increases of more than 5 metres according to the Intergovernmental Panel on Climate Change (IPCC). Were such a surge to happen again it would endanger cities such as New York, London and Mumbai.

Scientists have fretted about glacier melt for years, and over the West Antarctic Ice Sheet in particular since the 1970s. The Thwaites Glacier, the sheet's "weak underbelly", saw its grounding-line retreat 14 kilometres (8.7 miles) between 1992 and 2011, for example, and some fear nothing can be done to save it from collapse—irretrievable loss. Deep valleys in Greenland's ice sheet may mean it melts much faster than once expected, hastening sea-level rise and disrupting ocean-circulation patterns. And shrinking glaciers in mountainous parts of the world, such as California and northern India, mean farmers lack meltwater for their crops. Studying observations over four decades, the IPCC is confident that glaciers, alongside the thermal expansion of warmer water in the ocean, explain 75% of the rise in sea level in the past century.

Appendix II:

Survey Questions:

Block 1-4: Questions on news articles

Q1.1 How would you rate the quality of your reading?

❍ Extremely dissatisfied (1)

❍ Somewhat dissatisfied (2)

❍ Neither satisfied nor dissatisfied (3)

❍ Somewhat satisfied (4)

❍ Extremely satisfied (5)

Q1.2 How well do you feel you understand the graphics? (Visual)

How well do you feel you understand the information? (Text)

❍ Not well at all (1)

❍ Slightly well (2)

❍ Moderately well (3)

❍ Very well (4)

❍ Extremely well (5)

Q1.3 How often do you feel you stumbled while reading the story?

○ Always (1)

○ Most of the time (2)

○ About half the time (3)

○ Sometimes (4)

○ Never (5)

Q1.4 What useful information did you get from this story?

Q1.5 What is the main idea or proposition of this story?

Q1.6 How do you rate the value of the story?

○ Not at all important (1)

○ Slightly important (2)

○ Moderately important (3)

○ Very important (4)

○ Extremely important (5)

Q1.7 How informative do you think the graphics are? (Visual)

How informative do you think the story is? (Text)

○ Not at all (1)

○ Somewhat (2)

○ Quite a bit (3)

○ Moderately (4)

○ Very much (5)

Q1.8 Does this story make you want to read more related stories?

○ Definitely not (1)

○ Probably not (2)

○ Might or might not (3)

○ Probably yes (4)

○ Definitely yes (5)

Q1.9 Does this story make you want to get more involved in this issue?

○ Definitely not (1)

○ Probably not (2)

○ Might or might not (3)

○ Probably yes (4)

○ Definitely yes (5)

Q1.10 How do you evaluate the pace at which you read this article?

○ Extremely dissatisfied (1)

○ Somewhat dissatisfied (2)

○ Neither satisfied nor dissatisfied (3)

○ Somewhat satisfied (4)

○ Extremely satisfied (5)

Q1.11 How likely would you be to comment on this story?

○ Extremely unlikely (1)

○ Somewhat unlikely (2)

○ Neither likely nor unlikely (3)

○ Somewhat likely (4)

○ Extremely likely (5)

Q1.12 How likely would you be to recall this information later on?

○ Extremely unlikely (1)

○ Somewhat unlikely (2)

○ Neither likely nor unlikely (3)

○ Somewhat likely (4)

○ Extremely likely (5)

Q1.13 How likely would you be to share this story on your social media?

○ Extremely unlikely (1)

○ Somewhat unlikely (2)

○ Neither likely nor unlikely (3)

○ Somewhat likely (4)

○ Extremely likely (5)

Q1.14 How likely would you be to mention this story to your friends?

○ Extremely unlikely (1)

○ Somewhat unlikely (2)

○ Neither likely nor unlikely (3)

○ Somewhat likely (4)

○ Extremely likely (5)

Q1.15 Any comments, questions or concerns?

Block 5: Basic information

Q2.1 What's your age?

Q2.2 What's your gender/gender identity?

❍ Male (1)

❍ Female (2)

❍ Trans (3)

Q2.3 What's your ethnicity?

❍ White/Caucasian (1)

❍ African American (2)

❍ Hispanic (3)

❍ Asian (4)

❍ Native American (5)

❍ Pacific Islander (6)

❍ Other (7)

Q2.4 What's your nationality?

Q2.5 What's your native language?

Q2.6 How many years have you been learning English?

Q2.7 What are your best language test scores (TOEFL/IELTS)? (optional)

TOEFL (1)

IELTS (2)

Q2.8 What's your current education level?

○ Freshman (1)

○ Sophomore (2)

○ Junior (3)

○ Senior (4)

○ Graduate student (5)

○ Doctoral student (6)

○ Other, please specify. (7) _____

Q2.9 What's your major?

Block 6: News attitude and habit

Q3.1 Where did you get news yesterday? (Please choose what applies.)

❑ Mobile device (1)

❑ Computer (2)

❑ TV (3)

❑ Radio (4)

❑ Print (Newspaper, magazine) (5)

❑ Other, please specify. (6) _____

❑ N/A (7)

Q3.2 How often do you enjoy keeping up with news?

○ Never (1)

○ Sometimes (2)

○ About half the time (3)

○ Most of the time (4)

○ Always (5)