Fall 1-6-2023

“’NUFF SAID”: UNDERSTANDING COMPREHENSION PROCESSES AND PRODUCTS FOR READING TEXT AND NON-LINGUISTIC GRAPHIC NARRATIVES

Heather Ness-Maddox

ACCEPTANCE

This dissertation, "'NUFF SAID": UNDERSTANDING COMPREHENSION PROCESSES AND PRODUCTS FOR READING TEXT AND NON-LINGUISTIC GRAPHIC NARRATIVES, by HEATHER NESS-MADDOX, was prepared under the direction of the candidate's Dissertation Advisory Committee. It is accepted by the committee members in partial fulfillment of the requirements for the degree, Doctor of Philosophy, in the College of Education & Human Development, Georgia State University.

The Dissertation Advisory Committee and the student's Department Chairperson, as representatives of the faculty, certify that this dissertation has met all standards of excellence and scholarship as determined by the faculty.

____
Sarah E. Carlson, Ph.D.
Committee Chair

____
Elizabeth Tighe, Ph.D.
Committee Member

____
Joseph P. Magliano, Ph.D.
Committee Member

____
Hongli Li, Ph.D.
Committee Member

____
Date

____
Brendan Calandra, Ph.D.
Chairperson, Department of Learning Sciences

____
Paul A. Alberto, Ph.D.
Dean, College of Education &
Human Development

AUTHOR'S STATEMENT

By presenting this dissertation as a partial fulfillment of the requirements for the advanced degree from Georgia State University, I agree that the library of Georgia State University shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to quote, to copy from, or to publish this dissertation may be granted by the professor under whose direction it was written, by the College of Education & Human Development's Director of Graduate Studies, or by me. Such quoting, copying, or publishing must be solely for scholarly purposes and will not involve potential financial gain. It is understood that any copying from or publication of this dissertation which involves potential financial gain will not be allowed without my written permission.

____

HEATHER NESS-MADDOX

NOTICE TO BORROWERS

All dissertations deposited in the Georgia State University library must be used in accordance with the stipulations prescribed by the author in the preceding statement. The author of this dissertation is:

Heather Ness-Maddox
Department of Learning Sciences
College of Education & Human Development
Georgia State University

The director of this dissertation is:

Dr. Sarah E. Carlson
Department of Learning Sciences
College of Education & Human Development
Georgia State University
Atlanta, GA 30303

CURRICULUM VITAE

Heather Ness-Maddox

ADDRESS: 100 University Parkway, Suite 356
Macon, GA 31206

EDUCATION:

Ph.D.	2022	Georgia State University Educational Psychology
M.S.	2015	Clayton State University Clinical Psychology
B.S.	2013	Middle Georgia State University Psychology

PROFESSIONAL EXPERIENCE:

2022 – Present	Assistant Professor, Middle Georgia State University
2018 – 2022	Research Assistant, Georgia State University
2015 – 2018	Adjunct Instructor, Middle Georgia State University
Fall 2015	Research Assistant, Mercer University
2014 – 2015	Research Assistant, Clayton State University

PUBLICATIONS:

Ness-Maddox, H., Carlson, S.E., Dahl, A., Kennedy, P., Davison, M., Seipel, B., & Clinton-Lisell, V. (In print). *Emotion as a Factor of Causal Coherence in Text Comprehension.* Contemporary Educational Psychology

Ness-Maddox, H. (2022). Question, Discover, Apply, Disseminate: My Journey from Honors

Student to Educator. *Journal of the National Collegiate Honors Council, 23*(1), 85–87.

Ness-Maddox, H., Dahl, A., Reynolds, E., Carlson, S.E., Seipel, B., Clinton-Lisell, V., & Davison, M. (2020). Exploring the differences between college students' think-aloud responses for narrative and expository texts (No. 3825). EasyChair. https://easychair.org/publications/preprint/RmQ5

PRESENTATIONS:

Ness-Maddox, H., Dahl, A.C., Tadeo, M., Troemel, V., Berry, J., Carlson, S.E., & Magliano, J.P. (July 2022). *Emotional Inferences Due to Framing and Composition During Think-Aloud Task with Non-Linguistic Graphic Narratives.* Standard presentation at the annual meeting for the Society for Text and Discourse, Atlanta, Georgia. [International]

Ness-Maddox, H. (July 2022). Conceptualizing Narrative-Based Video Games as a Narrative Modality. Spoken presentation at the annual meeting for the International Society for the Empirical Study of Literature, Orlando, Florida. [International]

Ness-Maddox, H., Dahl, A., Reynolds, E., Carlson, S.E., Seipel, B., Clinton-Lisell, V., & Davison, M. (2020, July). Exploring the differences between college students' think-aloud responses for narrative and expository texts [Conference session]. Society for Text and Discourse Annual Meeting, Atlanta, GA, United States. https://easychair.org/smart-program/STD2020/2020-07-22.html

Ness, H., Carlson, S.E., Seipel, B., Clinton, V., Taylor, T., Bajpayee, S., Biancarosa, G., Davison, M. (July, 2019). *Item-Writing Methodology for MOCCA-C Reading Comprehension Assessment.* Poster presented at the annual meeting for the Society for Text and Discourse, New York, New York. [International]

Ness-Maddox, H., Dahl, A., Carlson, S.E., Seipel, B., Clinton-Lisell, V., & Davison, M. (July, 2021). A coding scheme to describe emotional dimensions in narrative and expository texts [Conference symposium]. Society for Text and Discourse Annual Meeting, Virtual.

Ness-Maddox, H., Carlson, S.E., Seipel, B., Clinton, V., Taylor, T., Bajpayee, S., Biancarosa, G., Davison, M. (October, 2020). How emotional information in a reading comprehension assessment affects examinees' response choices. Presentation at the Georgia Education Research Association Conference, Virtual. [State]

Ness, H., Carlson, S.E., Seipel, B., Clinton, V., Taylor, T., Bajpayee, S., Biancarosa, G., Davison, M. (October, 2019). *Item-Writing Methodology for MOCCA-C Reading Comprehension Assessment.* Presentation at the Georgia Education Research Association Conference, Macon, Georgia. [State]

“’NUFF SAID”: UNDERSTANDING COMPREHENSION PROCESSES AND PRODUCTS FOR READING TEXT AND NON-LINGUISTIC GRAPHIC NARRATIVES

by

HEATHER NESS-MADDOX

Under the Direction of Dr. Sarah E. Carlson

ABSTRACT

People encounter and comprehend narratives in a variety of modalities: text, graphic, film, audio, and others. Linguistic modalities (e.g., text, audio) require language comprehension while visual modalities (e.g., graphic, film) require visual comprehension and also language comprehension when text or audio is included. However, it is unknown whether readers engage in similar or different cognitive processes and construct similar or different comprehension products to the same extent for linguistic and non-linguistic narratives (i.e., with no text or audio). Thus, studies have not directly compared the in-the-moment (i.e., online) processes and post-reading (i.e., offline) products of comprehension across linguistic and non-linguistic narratives. A review of the current literature on graphic narrative comprehension is presented. The goal of this study was to explore the extent to which readers generate online cognitive processes and produce offline comprehension products post-reading across text and non-linguistic graphic narratives. A sample of 51 participants completed a think-aloud task with non-linguistc graphic and text versions of narratives to to assess readers’ online cognitive processes. A subsample of 48 participants also completed a recall task to assess their comprehension offline products (i.e., text / image base and situation model) post-reading. In addition, participants’ text print exposure and visual language fluency were measured to control for participants’ experience with both modalities. Overall, narrative modality had an effect on both participants’ comprehension processes and products.

Post-hoc analyses revealed that during the think-aloud task, participants generated more backward-oriented inferences (i.e., anaphoric, bridging) and generated more inferences about characters' emotions for non-linguistic graphic narratives. For text narratives, participants generated more forward-oriented inferences (i.e., predictions) and generated more statements about characters' goals. During the recall task, participants included more emotion inferences in their situation model representation for non-lingusitic graphic narratives but included more accurate story information for their text base representation for text narratives. These findings suggest that modality (i.e., linguistic or visual information) influences how readers process and comprehend narratives and are discussed in terms of theoretical, research, and practical implications.

KEYWORDS: narrative comprehension, visual comprehension, graphic narratives, text narratives

“’NUFF SAID”: UNDERSTANDING COMPREHENSION PROCESSES AND PRODUCTS FOR READING TEXT AND NON-LINGUISTIC GRAPHIC NARRATIVES

by

HEATHER NESS-MADDOX

A Dissertation

Presented in Partial Fulfillment of Requirements for the

Degree of

Doctor of Philosophy

in

Educational Psychology

in

Department of Learning Sciences

in

the College of Education & Human Development

Georgia State University

Atlanta, GA
2022

DEDICATION

For my wife Rachel Ness-Maddox, for your patience and constant support. Thank you for listening to my complaints and my frustrations, and for not allowing me to quit and open a comic book shop. At least not until the dissertation was finished.

For my mother, Mary Ness, for teaching me to read. Thank you, Mom. And for my father, Keith Ness, for buying me my first comic book. Thank you, Dad. I love you both.

ACKNOWLEDGMENTS

Thank you to the many individuals who helped on this study. Each of the committee members – Drs. Joseph Magliao, Elizabeth Tighe, and Hongli Li – provided advice, feedback, and encouragement throughout this study. Your support and guidance has been invaluable, and I look forward to future collaborations together.

Thank you to Meghan Tadeo, Amanda Dahl, and Virginia Troemel. As colleagues, your help coding data made an intimidating project much more doable. As friends, your support kept me focused and greatly encouraged. I could not have done this alone.

Finally, many thanks to my advisor and committee chair, Dr. Sarah Carlson, for her mentorship and guidance on this project. I appreciate the time you dedicated helping me craft this work.

Contents

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION TO THE STUDY

Reading is a frequent activity of daily life. Colloquially speaking, people not only read text, but they also “read” images in comics and graphic novels, “read” music, “read” a map, “read” another person’s face, and even “read” a room. How then is the act of “reading” defined? The Simple View of Reading (Hoover & Gough, 1990) defines *reading text* as a combination of decoding (i.e., interpreting written symbols as words) and language comprehension (i.e., connecting the written words to linguistic meaning). However, as mentioned, reading can also involve many other modalities, including graphic narratives, such as comic books and graphic novels. For instance, *reading graphic narratives* involves reading the images by “decoding” drawn images (i.e., visual symbols) and connecting those images to semantic meaning for visual comprehension. Indeed, whether someone reads a text or a drawn image, the reader still decodes (i.e., extracts or takes in information from the narrative) symbolic information from the narrative material and then interprets the symbolic information as meaningful content through linguistic or visual comprehension. What is less understood is whether information presented through linguistic or visual modalities influences how a reader understands a story. That is, once a reader decodes linguistic or visual information, does that reader then make the same connections to their background knowledge for both a text and graphic narrative? In order to understand comprehension across linguistic and visual modalities, this study compared reading across text and non-linguistic graphic narratives.

There are many theories of reading comprehension, and in this study, I draw upon four theories as a foundation to my research. First, I draw upon the Construction-Integration (CI) Model (Kintsch, 1988; 1998), which, although originally focused on reading comprehension

from text, in general, this theory has also been applied to the comprehension of other modalities (e.g., graphic narratives, art; Magliano et al., 2018; Millis & Larson, 2018). In order to comprehend any type of narrative modality, readers engage in cognitive *processes* to develop comprehension *products*. Cognitive processes (e.g., inferences, predictions, associations) are generated and updated by readers *during* reading or “online” (Graesser et al., 1994; Kintsch, 1988; 1998) in order to build mental representations of the narrative. That is, these processes are how readers develop their representation of what they understand. The *product*, or mental representation, thus becomes the reader’s understanding or memory of the narrative after reading or “offline” (Graesser et al., 1994). Generally speaking, according to the CI Model (Kintsch, 1988; 1998; Kintsch & Van Dijk, 1978), readers, thus, construct three levels of mental representation (i.e., products) for comprehension in memory: the surface structure, the text base, and the situation model. By accurately decoding information, readers form the *surface structure* representation of the visual information in memory. Memory of the surface structure is fleeing (i.e., 1-2 seconds), and therefore, the current study focuses on the next two levels of mental representation. By connecting decoded information to semantic meaning in the reader’s knowledge base, the reader forms the *text base* representation of the explicit semantic content in memory. Finally, by integrating semantic meaning from the discourse with prior knowledge and generating inferences to form implicit connections, the reader constructs a *situation model* representation in memory of the overall meaning.

In addition to the CI model (Kintsch, 1998; Kintsch & Van Dijk, 1978), and more recently, other theories have been developed to complement and extend the CI model to provide a deeper understanding of comprehension across modalities. Magliano et al. (2013), for instance,

developed a framework to understand what they characterize as the front-end and back-end processes that take place during narrative comprehension across text, graphic narratives, and film in order to develop a product of comprehension. They describe the front-end processes as the processes that readers use to obtain symbolic information from the narrative depending on the modality of the narrative. That is, readers take in information from the narrative either through text decoding or through viewing visual images. Magliano and colleagues describe universal back-end processes as processes that readers generate (e.g., inferences) regardless of modality to construct their situation model.

Previous research has shown that readers of both text and graphic narratives generate back-end processes such as inferences during reading to construct their situation models (Magliano et al., 2012; Magliano et al., 2016; McNamara & Magliano, 2009). However, while both text and graphic readers may generate inferences during reading, it is unclear if readers necessarily generate the same types of inferences while reading at the same points in a narrative whether a narrative is text or graphic. If readers do not generate similar inferences depending on the narrative modality, then differences in inferences during reading should lead to differences in situation model representations post-reading (i.e., offline). In addition, and connecting this work to the CI Model, Magliano et al. (2018) also explain that there is little research addressing whether construction of other comprehension products such as the text base and the situation model is the same across modalities (Cohn & Magliano, 2020; Kendeou et al., 2020; Coderre, 2020; Magliano et al., 2018).

The third theoretical framework that supports my study, the Inferential Language Comprehension (iLC) Framework (Kendeou et al., 2020) builds upon the concept of universal back-end process presented in the Magliano et al. (2013) framework. Specifically, the iLC Framework

proposes that readers of text and graphic narratives should generate the same types of inferences during comprehension processing for both text and graphic narratives to support the construction of the product or representation of comprehension. These types of inferences include anaphoric, bridging, elaboration, and emotion. That is, whether a reader reads a text narrative or a graphic narrative, the reader should generate these types of inferences during the act of reading. Previous research has supported this notion and identified that readers' general inferencing skills extends across modalities (e.g., Kendeou et al., 2008). However, what is unclear about this notion is whether readers generate different *amounts* of each inference type (i.e., anaphoric, bridging, elaboration, emotion) depending on modality.

That is, readers may generate different amounts of different types of inferences for text or graphic narratives due to the various text and non-linguistic differences between text and graphic narratives which emphasize different information (Magliano et al., 2013; Magliano et al., 2012). For example, readers generate bridging inferences during reading of both text and graphic narratives (e.g., Magliano et al., 2016; Hutson et al., 2018; McNamara & Magliano, 2009), it is less clear whether readers generate more bridging inferences for graphic narratives in order to infer the missing action between the sequential images. In addition, readers of graphic narratives may generate more emotional inferences because of the facial expressions of the characters presented visually. On the other hand, readers may also generate more inferences about characters' internal states during text reading since text can communicate information about characters' internal thoughts. While the literature supports that readers of both text and visual narratives generate similar types of inferences during reading (e.g., Magliano et al., 2016; Kendeou et al., 2008), future research can provide more detailed evidence regarding the differences between the nuances within each type of modality.

Finally, in addition to the theories that support how (i.e., the processes generated) readers develop their mental representations of linguistic and visual information, Cohn's (2020) Parallel Interfacing Narrative-Semantics (PINS) Model emphasizes the importance of narrative structure during readers' narrative comprehension. According to the PINS Model, readers process semantic content and narrative structure simultaneously while reading both text and graphic narratives. That is, while readers process the meaning of text words and visual images, they also use the structure of narrative to comprehend the story. For example, in the PINS model, Cohn notes that processing narrative structure at the same time as semantic content helps readers generate other processes such as predictions about what will happen next in the narrative. For example, when a reader is at the climax of the story (i.e., a part of the narrative structure that connects a goal with a resolution), they may generate predictions about the story's ending or overall theme based on knowledge about how stories tend to conclude. Furthermore, during reading, as readers access new information from the narrative, they confirm or update their predictions. To support these notions, Cohn and colleagues (2014) found that when reading non-linguistic graphic narratives, participants showed anticipatory neural activity before shifts in narrative events, indicating that participants were predicting a change in the story based on expectations of narrative structure. These findings are important because they suggest that not just narrative content but how content is organized (i.e., narrative structure) influences how readers process a narrative However, it is still unclear whether readers generate predictions to the same extent for both text and non-linguistic graphic narratives.

Two factors need to be considered with regards to readers' predictions made during reading of text and non-linguistic graphic narratives. First, readers can generate predictions during text and graphic narratives based on narrative structure if readers recognize the structure as they

read and organize narrative information into key event categories (i.e., starting place, disrupting event, creation of a goal, goal attempt, causal consequence, ending place). However, if a reader do not recognize the narrative structure, they may not process the information from the narrative necessary to generate a prediction.

Second, theorists for text, graphic narratives, and film have described key events that seem to appear across narratives that help readers to generate predictive inferences. Similar key events that theories of narrative structure tend to include are a starting place or an introduction to the setting, a beginning or initial event that starts the causal chain of events throughout the narrative, the formation of a goal, a goal attempt, and a causal consequence to the goal attempt (Johnson & Mandler, 1980; Stein & Glenn, 1979; Cohn, 2013; Cutting, 2016). According to the PINS Model, each event in a narrative is causally linked to the preceding and following event, which allows readers to generate predictions about future events during online processing of the narrative. However, one difference in narrative structure between text and graphic narratives may be an emphasis on characters' goals. That is, many text story grammars include goals as an integral part of a narrative (see Johnson & Mandler; 1980; Rumelhart, 1977; Stein & Glenn,1979 for more information about goal attempts as a key event in a narrative). However, visual narrative story grammars tend to not include goals and instead focus on actions. For example, Cohn's (2013) Visual Narrative Grammar and Cutting's (2016) film structure focus on actions that occur during a narrative rather than characters' goals driving the plot. Perhaps this difference in narrative structure between text and visual modalities stems from text narratives being able to explicitly express characters' internal states and goals, but viewers must infer characters' internal states and goals from purely visual modalities (Magliano et al., 2013; Magliano et al., 2012); thus, then hindering the reader from generating a prediction when reading graphic narratives.

Thus, whether or how differences in narrative structure across linguistic and visual modalities affects readers' predictions or the generation of other types of inferences of narratives is unexplored. In addition, studies have not investigated whether differences in narrative structure between text and graphic narratives encourage readers to generate different types of inferences used to develop their representations of text. For example, future research can explore whether this differences in narrative structure across text and graphic narratives leads to readers identifying characters' goals during reading and including such goals in their mental representations post-reading.

Thus, the current study is designed to address these identified areas of unexplored research. That is, considering these four theories of comprehension together creates a comprehensive theoretical foundation for the current study to investigate narrative comprehension across modalities. The CI Model (Kintsch, 1988; 1998; Kintsch & Van Dijk, 1978), Magliano et al.'s (2013) framework, the iLC Framework (Kendeou et al., 2020), and the PINS model (Cohn, 2020) all emphasize different elements of narrative comprehension across modalities but also complement each other. The CI Model emphasizes the construction of mental representations through knowledge activation, while Magliano et al.'s framework emphasizes types of front-end and back-end processes (e.g., inference generation) across modalities used to construct these representations. While the CI Model and Magliano et al.'s front- and back-end processing framework discusses general inference generation used during comprehension, the iLC Framework expands on the idea of inference generation by proposing specific types of inferences readers generate for both text and graphic narratives. Furthermore, the iLC Framework identifies generally that readers of text and visual narratives generate the same types of inferences during reading, it is unknown whether they generate each type of inference to the same extent across modalities.

Finally, the PINS model is unique from the other frameworks in that it emphasizes the importance of narrative structure for reader processing and emphasizes accessing narrative structure and semantic information, generating predictions, and updating mental representations as part of comprehension processing.

Thus, using these four theories of narrative comprehension together creates a foundation to investigate online comprehension processing and offline comprehension products across text and non-linguistic narratives. In the current study, I examine online processing and offline products across text and non-linguistic graphic narratives through the lens of these four theories. Since text narratives and non-linguistic, graphic narratives present different forms of information (i.e., linguistic, visual) and narrative structure; thus, requiring different front-end processes to extract the information to construct the surface structure mental representation, modality may also affect other online processes generated while reading. Modality may also affect the content of offline products (i.e., text base; situation model) constructed for a text and graphic narrative (Cohn, 2020; Kendeou et al., 2020; Coderre, 2020). Thus, the current study investigated the nuances in similarities and differences in online and offline comprehension text and non-linguistic graphic narrative reading comprehension.

CHAPTER II

COMPREHENSION OF GRAPHIC NARRATIVES: A REVIEW OF THE LITERATURE

As noted above, narratives are a ubiquitous part of life. People read books, watch movies and television series, observe 90-second narratives in commercials, listen to people's stories about their experiences, and play games with narrative features. A narrative can be communicated through text, oral language, film, graphic novels and comics, and other modalities. All of these types of narratives across modalities require cognitive effort and resources to comprehend. However, what is less understood is how narrative comprehension of visual information (e.g., static and dynamic images) presented in visual modalities compares to narrative comprehension of linguistic information (e.g., text, audio). There is a large body of research investigating comprehension of text and audio narratives (e.g., Levoy, 2003; Pellicer-Sánchez et al., 2020;); however, there is a less expansive but ever growing body of literature investigating comprehension of static visual narratives. Here I review several studies focusing specifically on the online processes (i.e., in the moment cognition) and offline products (i.e., post-reading memory) of graphic narrative comprehension to understand the state of the field more fully. Because comprehension processes and products of text narratives has been reviewed in other works (e.g., Clinton et al., 2020; Schwabe et al., 2022), I focus on the relatively sparse body of literature for comprehension of graphic narratives (i.e., narratives composed of sequential, static images).

Thus, the goal of this review of the literature is to discuss previous studies that investigated comprehension processes and products of graphic narratives. In addition, this review will identify explored and unexplored areas of research concerning the comprehension of graphic narratives that informed the research questions and hypotheses of the current study. The following literature review includes studies that have investigated the online processes and offline products

of graphic narrative comprehension identified from the CI Model (Kintsch, 1988, 1998; Kintsch & Van Dijk, 1978), the front-end/back-end framework (Magliano et al., 2013), the iLC Framework (Kendeou et al., 2020), and the PINS Model (Cohn, 2020) in order to discuss their findings and inform the research questions and hypotheses of this larger work.

It is first beneficial to clarify the scope and purpose of this review. The focus of this review is to understand the cognitive processes and products of graphic narrative comprehension. By doing so, I hope to show that comprehending visual narratives is a complex task it is own right. Prior assumptions and misconceptions of graphic narrative comprehension assume that comprehending graphic narratives is inherently easier than comprehending text. For example, The Visual Ease Assumption is a concept that assumes comprehending visual information is less difficult than comprehending text, since readers (i.e., seeing-capable readers) take in visual information through their visual senses and interpret visual scenes almost constantly in their daily life (Coderre, 2020). Reading text, however, must be explicitly taught and developed as a skill. Indeed, viewing an image of a dog, for example, and recognizing that visual information as a dog is easier than reading the word "dog," especially if the word is written in a language a reader cannot translate.

When images are arranged in a sequence to convey a story, however, the comparison of visual information to text becomes more complicated (Coderre, 2020). With a sequence of visual images, readers must not only take in visual information through their visual senses and recognize objects and characters within a scene, but they must also connect information across multiple images, infer actions and goals not explicitly shown, and decipher visual symbols to create a cohesive understanding of a story (Cohn, 2021; Cohn & Magliano, 2020; Magliano et al., 2013). These cognitive processes are necessary for graphic narrative comprehension and are aided by

background knowledge of and experience with comprehending visual modalities (Cohn, 2014). Many readers do not have background knowledge and experience reading graphic narratives, such as comic books and graphic novels. These readers may find comprehending a story written in text easier than comprehending the same story as a graphic narrative (e.g., Cook, 2017). Other readers have extensive experience with graphic narratives and may find a graphic narrative much easier to comprehend than a text narrative. Thus, ease of comprehension may be more related more to individual skill, background knowledge, and experience with a modality than necessarily the modality itself (Cohn, 2201) rather than merely assuming visual information is easier to comprehend (Coderre, 2020).

Therefore, the following review and corresponding study to follow assumes that graphic narrative comprehension is a complex cognitive task, related to a reader's experience with graphic narratives and reading skill. While visual comprehension may not be easier per se than linguistic comprehension, visual comprehension may differ from linguistic comprehension due to differences across modalities. Understanding the research on graphic narrative comprehension, thus, will illuminate the complexities of visual comprehension. The goal of the literature review is to understand the current body of literature that has investigated this complex task of graphic narrative comprehension.

Review of the Literature

The following review is informed by the online processes and offline products identified in the theoretical frameworks used for the foundation of this study. Thus, the purpose of this review is to discuss the studies which have investigated the processes and products identified by

the CI Model (Kintsch, 1988; 1998; Kintsch & Van Dijk, 1978), the front-end/back-end framework (Magliano et al., 2013), the iLC Framework (Kendeou et al., 2020), and the PINS Model (Cohn, 2020) for comprehension of graphic narratives and their findings.

Search Methods for this Review

To identify these studies, I used APA PsycInfo for my search of the literature. I only included peer-reviewed journals that were written in English. I did not set a time limit for publication date. For my article search, my search terms were based on the processes, products, and important concepts identified in the theoretical frameworks for graphic narrative comprehension. The online processes I specifically searched were access, anaphoric inferences, bridging inferences, elaborations, emotion, predictions, and updates (Kendeou et al., 2020; Cohn, 2020). The offline products I specifically searched were the text base and situation model representations (Kintsch, 1988; 1998; Magliano et al., 2013). I conducted four separate library searches using the following search terms. For the first search, I used the terms “visual OR graphic novels or comic books or comics OR images or pictures or visual information AND narrative AND comprehension AND inference OR emotion inference OR bridging inference OR elaborative inference OR anaphoric inference OR access OR prediction OR update.” For the second search, I used the terms “visual OR graphic novels or comic books or comics OR images or pictures or visual information AND Cognitive process* AND narrative structure.” For the third search, I used the terms, “visual OR graphic novels or comic books or comics OR images or pictures or visual information AND textbase AND memory.” Finally, for my fourth search, I included the search terms, “visual OR graphic novels or comic books or comics OR images or pictures or visual information AND situation model AND recall AND narrative.”

From the articles found through these searches, I eliminated articles unrelated to my purpose. For instance, I eliminated studies that investigated comprehension of text narratives or film narratives exclusively because these did not focus on graphic non-linguistic narratives. I did include studies that used text materials to compare text and graphic narrative comprehension in order to review the findings in terms of graphic narrative comprehension. I also eliminated studies that included only expository or informational materials or that focused on math skills or numeracy. I also eliminated studies that focused on narrative comprehension of participants with disorders, disability, or dysfunction, multiple language learners, non-native English speakers, and deaf or hard of hearing participants. Because the field of graphic narrative comprehension research is relatively new and growing, I did not eliminate studies based on age of the participants; that is, I retained studies whether they included child participants, college students, or older adults.

I then included articles based on the following criteria. I included articles primarily investigated comprehension of graphic narratives. I also included articles that were quantitative studies. Thus, I eliminated reviews, commentaries, and qualitative studies. In order to be included with my selected criteria, a study needed to include a measure of either the processes identified based on findings from any the four theoretical frameworks: anaphoric inferences, elaborative inferences, bridging inferences, emotion or internal state inferences, access or extracting information, predictions, or updates; as well as products or mental representations develop: the text base or the situation model. Studies were not found for all processes and products included as search terms. Thus, if the studies included any aspect I listed, they were included in the review. I also included studies that involved recall of stories or tested memory of stories because these are typical methods used to understand readers' mental representations of comprehension. Finally, I

included relevant studies that met the inclusion criteria that were cited in any of the included studies. Thus, using this inclusion criteria, I identified 22 articles.

The remainder of this chapter review the identified studies. Studies are separated broadly into two sections: studies that investigate comprehension *processes* of graphic narrative reading and studies that investigate comprehension *products* of graphic narrative reading. By doing so, I discuss the steps of processing visual content during graphic narrative reading in order to develop a mental representation of visual content after graphic narrative reading.

Comprehension Processes of Reading a Graphic Narrative

Imagine a reader picks up a graphic narrative, such as a comic book. Graphic narratives are a coherent sequence of visual images, with each image called a “panel” (McCloud, 1993). The reader’s eyes gaze at the first image or panel. By taking in visual information through their eyes, the reader is engaging in front-end processing (Magliano et al., 2013). Front-end processing involves taking in information from a narrative through the reader’s senses. Because different modalities present different forms of information (e.g., auditory, text, visual), readers engage in different types of front-end processing to match the form of information. Non-linguistic graphic narratives or graphic narratives without any text present only visual information, and readers engage in three types of front-end processing specific to visual information: scene gist, object/character, and sometimes motion processing. Scene gist processing involves the reader taking in the whole visual scene presented in a panel. Object processing involves the reader taking in visual information about individual objects or characters within the scene. Finally motion processing involves taking in visual information conveying motion. Graphic narratives traditionally include static images, and motion must be depicted with blurry images or motion lines, which are static visual symbols signifying motion in the static images. Through these three types

of front-end processing a reader takes in all the visual information presented in non-linguistic graphic narrative.

In viewing this visual information, the reader constructs the first type of mental representation: the surface structure (Kintsch, 1988; Magliano et al., 2013). The surface structure mental representation is the reader's limited and fleeting sensory memory of the visual information seen. This surface structure memory of the visual information shown (i.e., surface *form*) last only brief seconds in the readers' memory (McNamara & Magliano, 2009). This mental representation involves only the visual sensory information taken in through the eyes; the reader has not yet assigned any meaning to this visual information. As the reader takes in this visual information through the three types of front-end processing and constructs the surface structure representation, the reader must now think, that is, cognitively process, this information. Thus, the first step of processing this visual information is accessing the information and assigning meaning (Cohn, 2020).

Accessing the Information and Assigning Meaning

Front-end processing involves taking in the lines and color of a static, drawn image. Now the reader must assign meaning to the lines and color within the panel This process is referred to as access (Cohn, 2020). The reader views a panel (i.e., a single image within the visual sequence), takes in the drawn lines and color for the scene and objects or characters and recognizes what this visual information represents. For example, a reader may look at some of the lines within the panel in Figure 2.1 and think "booth" and look at the scene as a whole and think "diner." The reader has taken in visual information and assigned meaning to it. Access involves activating enough prior knowledge that the reader recognizes an object or a scene or recognizes visual symbols for what they represent.

A reader may also incorrectly assign meaning. For Figure 2.1, instead of thinking, "people in a dinner," the reader might think, "people on a bus." Errors such as these may stem from a reader's lack of experience viewing and recognizing visual information and symbols (Cohn, 2021). Thus, this example may depict how visual language fluency, which is how well a reader can interpret visual information and symbols in graphic narratives based on a reader's experience with graphic narratives, influences how accurately a reader can comprehend the images in a graphic narrative (Cohn, 2014). Much like how a person fluent in a language can quickly and accurately recognize and understand the meaning of spoken words of that language, a person high in visual language fluency can quickly and accurately recognize and interpret visual information in a graphic narrative.

Figure 2.1.

Graphic Narrative Panel Example

Note: Image is from *The Incredible Hulk* volume 2, issue 35 (Jones et al., 2002).

Some studies included here do not necessarily articulate their investigative focus as "access" per se, but they do measure behavioral evidence of participants attending to and interpreting visual information. For example, Foulsham et al. (2016) conducted eye-tracking studies to understand where comic readers place their attention while viewing panels from comic strips. They conducted two similar experiments. In their first experiment, participants viewed comic panels one at a time at their own pace. In their second experiment, participants viewed comic panels for a single strip or visual sequence shown all at once, two rows of three panels. In both eye-tracking experiments, participants viewed the comic sequences either in narrative order or scrambled order (i.e., panels showed in randomized order). Foulsham and colleagues found that

participants made fewer fixations in the narrative order condition. Panels in the scrambled order condition required more attention and, for the second experiment when participants saw the panels all at once, participants made more regressions or saccades back to previous information during the scrambled condition. Participants also fixated on different areas in the narrative depending on whether they saw panels in order or as scrambled. The researchers concluded that narrative context and structure (i.e., whether information presented in a coherent narrative order) affects viewers' attention for graphic narratives, that is, where participants look to take in visual information (i.e., access) the narrative. That is, the nature of the narrative context and structure guides graphic narrative readers to look for visual information that aids their comprehension of the narrative. In other words, how readers access visual information helps in their assignment of meaning to the information in the narrative.

In a follow-up study, Cohn and Foulsham (2020) created cropped or zoomed-in panels based on the fixation density findings with the prior eye-tracking data (Foulsham et al., 2016). That is, they cropped panels to show only areas where participants' had previously fixated in the narratives. In this study, Cohn and Foulsham had participants view comic strips without any text. All comic strips or sequences included six panels, and all sequences were shown either as "scene sequences" in which the panels were all full panels (i.e., original panels not cropped) or "zoom sequences" in which all panels showed only high density fixation areas of the panel. As participants viewed these comic strips, participants' neural activation was measured using electroencephalogram (EEG) to observe object identification (i.e., accessing information) and updating. Panels were displayed one at a time for approximately one second at a time. Participants' EEG showed greater processing cost for object identification within fixation panels than full panels

which provide more context and background information. That is, participants had more difficulty recognizing objects when seeing only a portion of the panel rather than the full panel. Cohn and Foulsham also found that sequences of only zoom panels require more updating than sequences of full panels. That is, while it was found that participants fixed on these areas in the prior study (Foulsham et al., 2016), Cohn and Foulsham found that zooming in on these areas and only being able to access limited visual information required participants to expend more cognitive effort than seeing and accessing that visual information in context of the whole panel. In this study, they compared participants neural activity when viewing full panels and when viewing only the areas of the panels that previous participants fixated when viewing. However, this study does not consider whether viewers benefit from contextual information (i.e., full panels) to identify objects or if simply viewing any zoomed in panel inhibits viewer processing. In their next study, Foulsham and Cohn (2021) include a control condition in which participants see areas of panels that previous participants did not fixate.

As a follow-up to this study, Foulsham and Cohn (2021) then conducted three experiments investigating processing of zoom-edited panels on fixation points to understand how the frame of a panel influences how well readers access information. In the first experiment, participants viewed comic sequences across three conditions: all panels as full panels, all panels as fixation zoom panels, or all panels as non-fixation zoom panels. These conditions were based on the eye-tracking findings from Foulsham et al. (2016) that identified fixation zoom panels as cropped to show only an area with high fixation density (e.g., areas where previous participants fixated, such as an important object in the panel), and non-fixation zoom panels as cropped to show only an area with low fixation density (e.g., areas where previous participants did not fixate, such as background area like the corner of a sidewalk). Participants viewed one panel at a

time on a computer and progressed through the visual sequences at their own pace by clicking a key. Foulsham and Cohn then measured and recorded participants' viewing times through jspsych across all the panels. They also asked participants to rate how easily sequences were to comprehend. Not surprisingly, participants rated non-fixation zoom sequences as harder to understand, since the non-fixation zoom panels tended to not show information necessary to understand the narrative. Foulsham and Cohn also found that fixation zoom panels were viewed longer than non-fixation zoom panels. They posited that fixation zoom panels have more relevant information to process (i.e., access) and interpret. Full panel sequences were rated as easiest to understand and viewed the longest. They concluded from this first experiment that although participants do fixate on specific areas of a panel, additional information in the visual scene of the panel (i.e., the full panel) provides narrative context which readers need to access to comprehend the narrative.

In their second experiment, Foulsham and Cohn (2021) sought to replicate their findings from their first experiment, but instead of manipulating all panels in a sequence, they only manipulated one panel in each sequence, the "initial" panel of the sequence. "Initial" panels in visual narratives are not necessarily the first panel in a visual sequence but do introduce the action of the sequence and has been identified as an integral panel for communicating narrative structure in visual narratives Cohn, 2013). Thus, the panel of these 6-panel sequences which introduces the main tension of the sequence was manipulated. Initial panels were manipulated so that they were either full panels, fixation zoom panels, non-fixation zoom panels, or incongruous fixation zoom panels (i.e., an initial panel from a different comic strip). In this study, Foulsham and Cohn measured participants' viewing times for the initial panels and the following "peak" panels (i.e., panels that depict the main action of an event; Cohn, 2013). Unlike in their first experiment,

Foulsham and Cohn found no viewing time difference between full panels and fixation zoom panels. That is, participants processed panels with only fixation areas and no background information just as quickly as panels with full background information. They also found longer viewing times for non-fixation zoom and incongruous zoom panels than full panels – for both peak and initial panels. That is, participants processed panels that did not show key information or were incongruous with the narrative longer than panels that showed important information and were coherent with the narrative. Participants rated sequences with full initial panels as easier to understand than sequences with fixation zoom panels, which in turn were easier to understand than both the non-fixation zoom and incongruous zoom panels, which supported findings their first experiment that readers more easily access information when the panels provide situational context.

In their third experiment, Foulsham and Cohn (2021) edited the initial panels to have an inset, in which the full panel is shown, but a box is drawn around either a fixation or non-fixation area and highlighted. This way, participants' attention was directed to either a fixation or non-fixation area, but background information was still available for participants to see. Again, Foulsham and Cohn measured viewing times for both the initial and peak panels. Viewing times were greatest for the non-fixation inset, then the fixation inset, and then the full panels. Although participants could see potentially helpful background information in the inset conditions, highlighting either the fixation or non-fixation areas caused longer viewing times, which the authors suggested as additional processing effort to access the information presented.

While Cohn and Foulsham's series of studies (Cohn & Foulsham, 2020; Foulsham & Cohn, 2021; Foulsham et al., 2016) focused on *where in the scene* comic readers focus their at-

tention of accessing information and how the presentation of important information guides readers to assign meaning, Cohn and Maher (2015) investigated viewers assessing visual information through *visual symbols*. Visual symbols represent a concept that would be difficult to explicitly depict. For example, motion lines (e.g., lines extending from a baseball flying through the air) visually represent the concept of motion since static images cannot explicitly depict motion. Cohn and Maher conducted two studies in which they investigated viewers' processing of the visual symbol of motion lines. In their first experiment, they measured participants' viewing times for comic strip panels shown one panel at a time. Participants saw a peak panel depicting an action with either normal motion lines, no lines, or anomalous or reversed lines. For example, a panel with normal motion lines may show a baseball player hitting a baseball with his bat. The ball is shown with motion lines coming from the swung bat to depict that the ball is moving away from the bat. A panel with no lines would not depict the motion lines. A panel with anomalous or reverse lines would show the lines on the other side of the ball, implying that the ball is moving toward the swung bat. Cohn and Maher measured shorter viewing times for normal lines, then no lines, then longest viewing times for reversed lines. In their second experiment, Cohn and Maher measured participants' neural activity using EEG while viewing the same panels. They found that the normal motion lines aid comprehension of events as found by participants' larger N400 event-related brain potentials (ERPs); that is, accessing visual symbolic information aids processing of the visual narrative sequence, and therefore, helps the reader comprehend the narrative.

Each of the above studies support the notion that readers extract visual information from graphic narratives (i.e., front-end processing; Magliano et al., 2013) to build their mental repre-

sentations of the graphic narrative. When confused or presented with unexpected visual information, viewers will search the panel or static image for information to help assign meaning of the narrative, unless the viewers recognizes that the panel does not contain relevant information (Foulsham et al., 2016). How visual information is presented (i.e., framed) to viewers also affects how viewers access the information (Cohn & Foulsham, 2020; Foulsham & Cohn, 2021). For example, a panel cropped to show only an object or area where a viewer might fixate their attention may provide enough visual information for the viewer to create a coherent representation of the information. Viewers also assess information through visual symbols like motion lines and assign semantic and narrative meaning if the viewer has the background knowledge to do so (Cohn & Maher, 2015).

By assigning meaning to explicit visual information in a graphic narrative panel, the reader constructs text base level of mental representation; that is, representation of the meaning of the explicit visual information shown in the graphic narrative. Traditionally this part of the representation is called the text base because it tends to refer to the explicit meaning of text (e.g., a word or a phrase). When discussing non-linguistic graphic narratives without text, a more accurate term would be "image base," in order to describe the explicit meaning of an image.

To comprehend a graphic narrative, the reader does not just view a single panel or image. The reader views a series of sequential images when, when cognitively connected, tell a coherent story (e.g., Magliano et al., 2016). As the reader views the sequence of visual images, accessing multiple pieces of visual information and assigning meaning to the visual information, the reader also generates other cognitive processes to connect these sequential images to their background knowledge in order construct the last level of mental representation: the situation model. The situation model is the reader's understanding and memory of the story. This understanding involves

the characters, their goals, the setting, the events that make up the story, and any other relevant information to comprehend the narrative.

To construct the situation model, the reader generates a variety of back-end processes. Unlike front-end processes, back-end processes do not depend on the form of information presented in a narrative (Magliano et al., 2013). Readers generate similar back-end processes, namely inferences, during graphic narrative comprehension. Researchers have identified many types of inferences, all of which serve to develop a situation model (Cohn, 2020; Kendeou et al., 2020). Next, I discuss research investigating the generation of anaphoric inferences, bridging inferences, emotion and internal state inferences, and elaborative, as well as predictive inferences during graphic narrative reading, used to develop their situation model of graphic narratives.

Anaphoric Inferences to Connect Objects and Characters

One way readers connect sequential images in order to develop a coherent situation model of a graphic narrative is through the generation of anaphoric inferences (Kendeou et al., 2020). Anaphoric inferences are represented when a reader recognizes an object, character, or place as the same object/character/place seen in previous panels. For example, when viewing the panel in Figure 2.1, a reader may think, "That's the same man I saw in the last panel opening a door. Now he's in a diner." The reader recognizes the character as the same character seen before and in doing so increases their coherent understanding of the event (Trabasso & Magliano, 1996; Trabasso & Suh, 1993).

Only one study met the inclusion criteria for this review that specifically investigated anaphoric inferences. Specifically, Coopmans and Cohn (2022) investigated viewers processing of anaphoric objects in comic strips without text. Participants saw six-panel sequences with each panel shown one at a time for approximately one second each. Either the second or third panels

of these visual sequences were cataphoric panels, that is, panels which show target objects for the first time in a sequence (e.g., first time the viewer sees a football in the visual sequence). These cataphoric panels were either full panels or cropped to zoom in on the target object. Anaphoric panels reshowed the target object later in the sequence and were also either full panel or zoomed in on the object. Coopmans and Cohn measured participants' neural activity through EEG as they viewed the sequences. They found that participants recognize objects they have previously seen and generate anaphoric inferences to connect the object currently being seen as a previously seen object. Without these anaphoric inferences, a reader may simply interpret an object as a new object (e.g., "I saw a football before and now I'm seeing another football."). By generating these connections to previous information, readers create a coherent mental representation of the graphic narrative (Trabasso & Suh 1993; Trabbasso & Magliano, 1996; Trabasso et al., 1989). While anaphoric inferences increase coherence of mental representations by connecting objects and characters, bridging inferences increase coherence by connection actions or events.

Bridging Inferences to Connect Events

Generating bridging inferences during graphic narrative comprehension often involve inferring about an action not explicitly shown on a current panel but that must have happened in the previous panel in order for one panel to coherently lead to the next panel (e.g., Magliano et al., 2016; Trabasso & Suh, 1993). For example, a panel may show a man outside a closed door. The next panel may show the same man standing in a new room, the previously seen door now open. The reader should generate the bridging inference between the two panels that the man opened the door and entered the new room. Without this inference, panels in a graphic narrative

are simply a series of separate images, rather than a coherent, ordered sequence of connected images communicating a whole story.

Perhaps because bridging inferences are so integral to building coherent mental representations of a narrative, researchers have conducted more studies focused on bridging inferences than other types of inferences generated during graphic narrative reading. For example, Cohn and Wittenberg (2015) investigated readers' generation of bridging inferences in two experiments. In their first experiment, they measured how long participants viewed panels as they read six-panel comic strips one panel at a time on a computer. Comic strips were in one of the following four conditions. Panels of the comic strips were presented in either coherent or scrambled (i.e., randomized) order. Narratives were also presented either with a peak panel (i.e., panel that shows an integral action or event of the narrative; Cohn, 2013) depicted or with an "action star" replacing the peak panel. Like motion lines, an action star is a visual symbol that replaces an action, often violent, in comics (e.g., an action star with the word POW! covering Batman's fist hitting a villain's face). Cohn and Wittenburg compared participants' viewing times on the panel after either the peak panel or the action star panel, that is, they measured viewing times on the panel depicting the end or resolution of the event. They found that the participants viewed panels after the action star in the coherent sequences the longest of the four conditions, implying that they were generating bridging inferences for what the action star was replacing to achieve the end scene.

In Cohn and Wittenburg's (2015) second experiment, comic strips were again in one of four conditions, this time with the original peak panel, the peak panel replaced with an action star, the peak panel replaced with a blank panel (i.e., a white screen), or the peak panel replaced with a peak panel from another narrative that was not coherent with the current narrative. Cohn

and Wittenburg measured how long participant viewed both the peak panel and the panel after the peak panel when seeing the panels of the narrative one at a time on a computer screen. While there was no difference in viewing time for the panel following the action star or the blank panel, participants did view the action stars for significantly less time than the blank panels. This implies that participants interpret visual symbols like an action star as representing information in a graphic narrative, unlike a controlled blank panel that depicts no visual information. The authors concluded that whether participants saw the action star or the blank panel, they generated a bridging inference to infer the missing action. That is, when the narrative introduced a coherence gap either through an action star or a blank panel, participants generate a bridging inference to complete the coherence gap (e.g., Magliano et al., 2016). In the next study, Manfredi et al. (2017) test whether readers generate a bridging inference whether a panel depicts visual information congruent or incongruent to the narrative.

Manfredi et al. (2017) built upon the findings from Cohn and Wittenburg's (2015) two experiments by measuring neural activity through EEG rather than viewing times when participants viewed comic panels. Participants saw comic strips without text, one panel at a time. In their first experiment, peak panels were replaced with an action star with onomatopoeia (e.g., bang!), descriptive text (e.g., impact!), anomalous onomatopoeia (e.g., smooch!), or symbols for swear words, called grawlixes (e.g., @#$%!). They observed neural activity (i.e., N400) that indicates participants generated bridging inferences for all the panels except for the grawlixes. In their second experiment, they replaced the grawlixes panels (i.e., the ones which did not elicit N400 activity) with anomalous descriptive text (e.g., kiss!). They found that anomalous and descriptive panels were more difficult to access because they were incongruent with the rest of the visual information in the narrative. In which case, the researchers observed less neural activity

when participants saw anomalous and descriptive panels, indicating they did not generate a bridging inference. They found more neural activity indicative of bridging inferences for congruent than incongruent panels. That is, viewers do generate bridging inferences when presented with visual symbols, if the context of visual symbol supports the inference.

In order to compare readers' generation of bridging inferences for visual symbols with processing explicit events, Cohn (2021) conducted a follow up of the previous study by Cohn and Wittenburg (2015) with action stars replacing depicted events. Cohn showed visual sequences with either the peak panel depicting an explicit event, an action star, or "noise" (i.e., scrambled lines). Rather than measuring participants' viewing times as in Cohn and Wittenburg's (2015), Cohn measured neural activity through EEG. He observed greater N400s for noise panels, then for action star, then for explicit events. This implies that participants had difficulty integrating noise panels into their representations, and that they integrated action stars into their mental representation less well than explicit events. He also observed P600 activity for action star panels. This implies that that action stars were integrated into participants' mental representations to stand for the action not shown. That is, participants interpret the action star as a visual symbol for an action and generate a bridging inference for the action replaced by the action star, but this process requires additional cognitive effort than when an action is explicitly shown. Additionally, Cohn posited that participants do not interpret noise panels as communicating story information. Depending on what visual information or symbols are shown, participants may not generate a bridging inference, not unless the narrative context supports the inference. In the next study, Cohn and Kutas (2015) investigate readers' bridging inferences when presented with visual information through facial expressions.

Cohn and Kutas (2015) move away from action stars as visual symbols and instead investigated whether facial expressions of characters could prompt readers to generate a bridging inference. They created comic strip sequences of four panels. A sequence was in one of four conditions: (1) an impoverished sequence in which the third panel contained an expressionless character looking at something outside the panel; (2) an implied sequence in which the third panel contained the same character but with a subtle emotional expression and either an exclamation point or question mark in a dialogue bubble above the character's head; (3) an expected sequence in which the third panel depicted the expected action in the sequence; and (4) an explicit condition in which the viewers saw an unexpected action in the third panel, but the action still made the sequence coherent. Cohn and Kutas measured participants' neural activity through EEG while viewing the third panels of the sequences and found larger neural positivity for the implied than impoverished panels. The researchers interpreted this finding as that the emotional expression and dialogue bubble was sufficient visual information for viewers to infer that an action was taking place outside of the panel and must have generated a bridging inference.

These studies, thus far, support the generation of bridging inferences during graphic narrative comprehension. Although these findings are important for understanding the nature of the generation of bridging inferences during graphic narrative comprehension, it is worth understanding if there are other components that help or hinder these to take place during comprehension. For instance, generating bridging inferences requires cognitive effort and resources and may place demands on the readers' working memory. To investigate the cognitive demands, such as working memory, on bridging inferences, Magliano et al. (2016) conducted three experiments to compare how demands on linguistic working memory and visuospatial working memory affects generation of bridging inferences while viewing non-linguistic picture stories.

All three experiments included six non-text picture stories which the participants read. Each story contained multiple action or event sequences. Event sequences contained a beginning state image, a bridging event image, and an end state image. Magliano and colleagues tested two competing hypotheses: (1) the visuospatial primacy hypothesis and the shared systems hypothesis, which states that visuospatial working memory load should impair bridging inferences, but a verbal working memory load should not; and (2) the shared systems hypothesis, both load types (visuospatial and verbal) should impair bridging inferences.

Their first experiment included both a pilot study and experiment. For both the pilot and the experiment, event sequences in the picture stories were presented with the bridging event image either present or absent. In the pilot study, participants typed what they understood about the event after seeing the end state image. This was to ensure that for bridging absent conditions, participants did infer the missing bridging event. In the experiment, participants saw the picture stories, and the researchers measured how long participants viewed the end state image after either the absent or present bridging event. They found longer viewing times for the end state panel in the bridging event absent condition. These findings from both the pilot study and the experiment support that viewers generate bridging inferences for non-linguistic picture stories.

In Magliano et al.'s (2016) second experiment, participants again saw the same non-linguistic picture stories with either bridging-present or bridging-absent event sequences, but this time, the researchers gave participants either a visuospatial or verbal load task (e.g., participants either had to remember a sequence of seven words or remember the spatial location of a sequence of seven dots), or participants received no load task. After presented with the load task, participants then saw a target event sequence in the picture story (either bridging-present or bridging-absent conditions). After the target event, participants were prompted to remember the

load sequence after viewing the end state image. As in their first experiment, Magliano and colleagues measured viewing times for the end state image. They found longer viewing times for bridging-absent conditions, as in their first experiment. They also found longer viewing times for load-absent conditions, suggesting that working memory load impairs viewers' generation of bridging inferences. The longest viewing times were for bridging-absent and load-absent conditions. They found no significant different between load type conditions, supporting the shared systems hypothesis.

In their third experiment, Magliano and colleagues (2016) replicated the results of their second experiment, but also included an articulatory suppression task (e.g., participants either saying "the" or clicking mouse while viewing images or no suppression task). Overall, results of their third experiment replicated the results of their second experiment. Additionally, they found no significant findings to support that sub-vocalization plays a part in generation of bridging inferences. Overall, findings across the three experiments support that viewers of non-linguistic graphic narratives do generate bridging inferences and, more specifically, that taxing both linguistic and visuospatial working memory affected viewers ability to generate bridging inferences during these non-linguistic graphic narratives.

If generating bridging inferences requires working memory cognitive resources, then generating bridging inferences may also affect memory of the graphic narrative. Next, Magliano et al. (2017) tested the memory distortion hypothesis, the idea that generating inferences during reading visual stories leads to poor memory for explicit story content. That is, generation of bridging inferences to construct the situation model during graphic narrative processing may affect the accuracy of the image base level of representation. They conducted two experiments to

test this hypothesis. In their first experiment, participants viewed picture stories without text content. As in Magliano et al. (2016), every picture story included multiple events which include a beginning, bridging, and end state panels. Events were in one of four conditions: complete event, missing beginning panel, missing bridge panel, or missing end state panel. Participants then completed a recognition task of panels viewed and not previously viewed. Magliano and colleagues measured how long participants' viewed individual panels reading the picture stories. Magliano and colleagues found longer viewing times in the condition when the bridge panel was missing. They observed more false alarms in the recognition task for beginning and bridge panels than end state panels. That is, participants inaccurately indicated that they remembered seeing bridging events they inferred but were not explicitly shown. The researchers replicated these findings in experiment two. Again, they found participants displayed better memory for end states than beginning or bridging panels, which supports the memory distortion hypothesis. That is, while generating bridging inferences are necessary to create a coherent representation of the narrative, generating bridging inferences of not-depicted events distorts a readers' memory for the visual information depicted. However, this study does not indicate whether this memory distortion for explicit information also occurs when readers generate bridging inferences during text reading.

While Magliano et al. (2016) and Magliano et al. (2017) both support that bridging inferences require cognitive resources from readers' working memory, they did not investigate what readers do in order to generate bridging inferences when their working memory is taxed. Hutson and colleagues (2018) investigate how readers behave when needing to generate a bridging inference while reading a graphic narratives. Hutson and colleagues use materials from Magliano et al. (2016) to test both the computational load hypothesis and the visual search hypothesis in an eye-tracking study. According to the computation load hypothesis, viewers of graphic narratives

have longer viewing times for end state images (e.g., Magliano et al., 2016) when required to generate a bridging inference due to *longer* fixations, suggesting that generating a bridging inference taxes working memory resources. The visual search hypothesis attributes longer viewing times to *more* fixations, suggesting that needing to generate a bridging inferences prompts readers to search the image for information to support the inference. The researchers manipulated picture stories by removing event panels, creating a condition in which participants would generate a bridging inference. They then recorded and compared participants' eye-movements on the panel after the present or removed event panel. If participants in the event absent condition (i.e., the condition that prompted a bridging inference) engaged in longer eye fixations, this would support the computational load hypothesis. If participants in the event absent condition engaged in more fixations, this would support the Visual Search hypotheses. They found that participants did engage in more eye fixations, which supported the visual search hypothesis. The researchers then asked a second set of participants to identify inference-relevant regions of the panels in the picture stories. These identified regions predicted the fixations of other viewers. That is, participants were able to verbally identify important visual information that support the generation of bridging inferences. This study supports that viewers of non-linguistic graphic narratives do generate bridging inferences, and that participants' active mental representations of the story influence their eye movements when generating bridging inferences. While generating a bridging inference may not tax working memory such that readers fixate longer while they process and generate the inference, needing to generate a bridging inference does prompt readers to search for visual information to support a bridging inference.

Hutson et al. (2018) show that when needing to generate a bridging inference to fill a causal gap in graphic narratives, readers search the visual images. Magliano et al. (2016) and

Magliano et al. (2017) also found that bridging inferences require working memory resources and distort memory of the explicit information in the images. These studies do not, however, investigate whether these findings are similar for text or how readers generate bridging inferences when taking in both text and visual information through a graphic narrative. Huff et al. (2020) used the picture stories from the prior studies (e.g., Magliano et al., 2016) to investigate how readers generate bridging information when presented with both text and visual information when reading a graphic narrative, what they term "cross-codal" integration. Across three experiments, they presented non-linguistic picture stories and manipulated the stories so that bridging event information was either present or absent and recorded how long participants viewed end state pictures. In their first experiment, if the bridging event information was present, it was presented as either an image or as a text sentence. Viewing times were longer in the absent condition, replicating Magliano et al.'s (2016) findings. Additionally, viewing times were longer in the text condition than in the picture condition which extended previous findings. This finding suggests that switching from visual information to text information increases the difficulty of integrating bridging information.

In their second experiment, Huff et al. (2020) replicated the procedures from their first experiment except they replaced the bridging-present text condition with a blank panel. Viewing times were longer in the blank panel condition than the bridging absent or picture condition. In their third experiment, Huff et al. directly compared picture, text, and blank conditions and did not include a bridging absent condition. Viewing times were longest for the blank condition followed by the text condition. Viewing times were shortest for the picture condition. Overall, these findings suggest that integrating information (i.e., generating a bridging inference) across multi-

ple modalities is more difficult than integrating information from one modality. Huff and colleagues did not include a comparative condition in which the story was mainly text and then bridging information was replaced with an image or a blank panel, so it would be misleading to conclude that integrating text information is harder than integrating visual information. Rather, integrating more than one source of information requires more cognitive resources than integrating one source of information.

Emotion and Internal State Inferences to Understand Characters

In addition to the above inferences, reader may generate inferences about a character's emotions or internal state during graphic narrative comprehension (Kendeou et al., 2020). For example, referring back to the previously presented panel in Figure 2.1, the reader might think, "The girl's eyebrows are raised, and her eyes are big. She must be scared of something." In this case, the reader would be generating an inference about the character's emotions based on understanding how the girl's facial information gleaned through object processing (Magliano et al., 2013). On the other hand, the reader might think, "The girl is thinking about what to order." In this case, the reader would be generating an inference about the character's internal state based on their prior knowledge of what people think about when in a diner. Only two studies met inclusion criteria for this review that investigated emotion or internal state inferences.

First, Gnepp (1983) investigated children's generation of emotion inferences for a graphic narrative based not only on characters' facial expressions but also on other emotional clues in a panel. Gnepp's sample included preschool, first grade, and sixth grade students. Gnepp investigated children's ability to make emotional inferences when graphic narratives present conflicting visual clues. Thus, the pictures used in the study depicted a situation (either positive or negative emotional valence) with a child in the situation showing an emotional expression either

congruent or incongruent to the situation. For example, one image may show the scene of a child's birthday party, with colorful party decorations, a pile of presents, and a large cake. In general, such a scene would activate associations with positive emotional valence and would provide a congruent situation because all the objects listed have positive valence associations. However, in an incongruent situation, a scene may show a child may siting on a stool in the corner of the room, arms crossed, and frowning. While the overall scene seems positive, the depiction of the child sitting alone in the corner of the room would prompt an association of negative emotions incongruent with the rest of the positive scene.

In Gnepp's (1983) study, children were assigned to one of four conditions: 1) the picture provided emotional cues either through scene or facial expression, 2and 3) the picture provided emotional cues through both the scene and facial expression and these cues were either congruent or incongruent, and 4) the experimenter provided verbal narrative context along with the picture (e.g., "Timmy just broke his bicycle."). For each story, children were asked to state out loud what they thought the child in the picture was feeling. Overall, Gnepp found that children tended to make emotional inferences based on the facial expression of the depicted child within the scene, although this tendency got weaker with age. That is, sixth grade students generated emotional inferences based on half being from facial expressions and half from the scene. Furthermore, children who heard the verbal narrative context along with seeing the picture were also more likely to base emotional inferences on the situation than children who saw the picture alone. This study suggests that as children age from pre-school to sixth grade, they begin to integrate conflicting visual information to generate emotional inferences about characters and a visual scene. That is, narrative context, not only facial expressions, can provide additional information to generate emotional inferences.

These findings are similar to Foulsham and Cohn's (Foulsham et al., 2016; Cohn & Foulsham, 2020; Foulsham & Cohn, 2021) work with zoomed in panels and fixation and non-fixation panels. Foulsham and Cohn's set of studies suggest that how visual information is presented (i.e., framed) to viewers also affects how viewers access the information (Cohn & Foulsham, 2020; Foulsham & Cohn, 2021). Additionally, Gnepp (1983) found that older children incorporate cues from background information to generate nuanced emotion inferences. Together, these findings suggest that while graphic narrative readers do fixate and pay attention to specific objects and characters within a panel or visual scene, background information or scene information influences the reader's understanding of the narrative. In Gnepp's study, older children based their emotion inferences on scene information as well as facial expressions. In the next study, Lewis and colleagues (1994) move away from emotion inferences and investigate children's ability to generate inferences about characters' internal states.

Lewis et al. (1994) investigated preschoolers' ability to generate inferences about picture book characters' internal states (i.e., thoughts, cognition, beliefs). Materials used was a picture book about a character who places her cat somewhere, but the cat changes location while the character is not looking. Readers should generate the internal state inference that the character has a false belief about where her cat is. Lewis and colleagues conducted five different experiments with preschoolers from approximately 3 years to 5 years old. In their first experiment, children were four to five years old. Half the children had the story read to them once and half narrated the story back to the experimenter after hearing the story. At a certain point in the story, the experimenter asked the participants where the character would look for her cat. Children would either respond that the character would look for the cat where she placed her cat (i.e., an accurate inference about the character's internal state and false belief) or that the character would

look for the cat where the cat is hiding (i.e., an incorrect inference about the character's internal state). The children tended to answer the question correctly if they narrated the story to the experimenter. This finding suggests that either telling the story back themselves helped the children to generate an accurate internal state inference about the character's false belief or that being exposed to the story information twice helped the children generate an accurate internal state inference. Consequently, in their next experiment, Lewis and colleagues (1994) tested whether being exposed to the story twice influenced children's internal state inferences.

In the second experiment, in order to test whether engaging with the story information twice influenced children's false belief inferences, a third condition was added in which the experimenter read the story twice. Again, more children who read the story back tended to answer the internal state (i.e., false belief) question correctly, followed by those who heard the story twice. This finding suggests that children aged 4 to 5 years old are able to generate inferences about character's internal states, specifically characters' false beliefs, and that actively interpreting the visual information and telling the story themselves aids their understanding of characters' internal states. Lewis and colleagues then tested whether children younger than 4 to 5 years old are also able to generate these inferences.

In the third experiment, the study involved younger children under 3 years and 9 months old and replicated the procedures in the second experiment. This experiment was conducted to investigate whether younger children were developmentally capable of inferring a character's internal state. This group of children were able to answer the internal state question if they remembered the events leading up to the character having a false belief during their retelling of the narrative back to the experimenter. In their fourth experiment , the researchers added an additional picture to the beginning of the story to show the location of the cat, and one-third of the children

were asked “where does she think her cat is” instead of “where will she look for her cat?” Asking children were the character thinks her cat is lead to fewer correct answers than asking about the character’s future action. Their fifth experiment replicated these findings, but this time half the children were asked where the character would look, and half were asked what the character was thinking. As in experiment four, more children were able to correctly answer the question about the character’s actions than about the character’s thoughts (i.e., internal state).

Overall, young children are able to generate constrained inferences about character’s internal states from graphic narratives, if an adult reader is telling the story to them at least once and if they actively construct their mental representation of the narrative with the experimenters’ help. Additionally, children must understand the narrative events that influence characters’ internal states, and they may better understand characters’ internal states in terms of their future actions as opposed to their thoughts.

Gnepp (1983) and Lewis et al. (1994) provide findings regarding children’s development of emotion and internal state inferences, respectively. However, no studies which met inclusion criteria in this review has examined the extent to which older children or adult readers generate emotion or internal state inferences while reading graphic narratives. The next set of studies investigates other inferences generated during graphic narrative reading for young children as well as older adults.

Other Inferences Generated During Graphic Narrative Reading

In addition to anaphoric, bridging, emotion, and internal state inferences, readers also generate other inferences once a reader assigns meaning to an image during graphic narrative comprehension (Kendeou et al., 2020). Readers should generate elaborations when accessing visual information activates prior knowledge (Cohn, 2020; Kintsch, 1988; 1998). For example, a

reader may view the panel presented previously in Figure 2.1 and think, “A man in a diner. Diners are where people eat, so the man is hungry.” Elaborations expound on the explicit information depicted in an image. Currently no prior research has examined readers generating elaborations while viewing non-linguistic graphic narratives. While none of the studies that met inclusion criteria investigated the generation of elaborations explicitly, I include studies that measure inferences broadly during graphic narrative comprehension. The following studies measure participants’ inferences generated during graphic narrative reading.

First, Tompkins et al. (2012) investigated the types of inferences preschool children generate during reading of non-text picture books, whether these inferences related to their comprehension of story books read to them, and, if so, whether these inferences predicted their story comprehension. Tompkins et al. (2012) presented a non-text picture book to 4- to 5- year- and asked the children to narrate the story while viewing the pictures. The researchers recorded their verbal responses and later coded for the number of inferences children generated while narrating the non-text picture book. In addition to measuring the total amount of inferences generated, these inferences were also coded for type or category. Categories of inferences were derived from Kendeou et al. (2008) and McGinnis et al. (2008) and included goals, actions, causal antecedents, causal consequences, activities, character states, dialogue, emotions, place, and objects. Thus, children had a general inference score as well as a score for each category. Children generated inferences about character activity the most and then inferences about character state second most. Children generated dialogue and emotion inferences more than goal and action inferences, and they generated causal consequence or causal antecedent inferences the least frequently.

After narrating the non-text picture book, children were read a children’s book with text and asked comprehension questions developed and validated for the story by Paris and Paris

(2003). These questions assessed the children's story comprehension, which Tompkins and colleagues (2012) defined as comprehension of narrative events within a story. A total of 10 comprehension questions were asked, seven questions at pre-determined points during the story as it was read and three questions at the end of the story. Total inference scores from the non-text picture story were significantly correlated with total text story comprehension scores. Specifically, goal, action, and character state inferences were significantly correlated with story comprehension. While total inference scores from the picture story were related to the text (listening) story comprehension scores, children's general inferences scores from the non-text picture book did not predict their text story comprehension when the researchers controlled for age and expressive vocabulary. Scores for goal, action, and character state inferences did significantly predict story comprehension after controlling for age and expressive vocabulary. The researchers concluded that some, but not all, inferences generated during the picture story were uniquely predictive of text story comprehension. Inferences that related specifically to the structure of the narrative (i.e., goals, actions, characters) predicted overall story comprehension. This finding supports that processing narrative structure aids the processing of narrative content. The next study provides evidence that organizing narrative information based on narrative structure may depend on age and cognitive development.

Poulsen et al. (1979) compared the think-aloud and recall of 4-year-old and 6-year-old participants reading picture stories without text. Picture stories were shown either in a normal condition (i.e., intended narrative sequence) or scrambled (i.e., random order). Children described each picture as they viewed them and then recalled the stories. In the normal condition, both age groups described the pictures as a story as they viewed the pictures. That is, they recognized that the individual pictures together told a coherent narrative. In the scrambled condition,

the 6-year-olds attempted to connect the randomized pictures as a cohesive story, but the 4-year-olds tended to label or describe the pictures as individual pictures. That is, the 4-year-old participants did not generate inferences in order to try to understand the scrambled pictures as a coherent narrative. The 6-year-olds generated inferences to make sense of the scrambled pictures and also generated inferences about characters' emotion and internal states. Tompkins et al. (2012) and Poulsen et al.'s (1979) studies suggest that children as young as 4-years-old can understand visual sequences of images as a coherent graphic narrative and generate inferences related to the narrative structure.

While Tompkins et al. (2012) and Poulsen et al. (1979) investigated visual inferences of children, Ribeiro et al. (2010) investigated visual inferences of older adults. Ribeiro and colleagues recruited 45 typically developing elderly (i.e., 61-82 years) Brazilian adults. They divided their participant sample into three groups based on years of education: 1-4 years, 5-8 years, and 9+ years. The adults were shown four pictures that had been selected from a pool of pictures chosen by 15 speech therapists. Each picture depicted a visual narrative scene (e.g., a woman peering through a car window at keys in the ignition). The pictures were described in terms of visual complexity (i.e., the number of objects in each picture) and inferential complexity (i.e., the estimated amount of processing required to understand the scene). The four pictures chosen for this study were all inferentially simple but separated into two groups: visually simple and visually complex. The adults were told to view each picture one at a time and describe what was happening in each one. Responses were coded as essential information (i.e., necessary inferences for comprehension), accessory information (i.e., pertinent but not necessary for comprehension inferences), non-pertinent information, and incoherent information. Overall, groups with fewer years of education generated less essential information inferences than those with more years of

education, for both visually simple and visually complex pictures. Only the group with 1-4 years of education significantly differed in the number of inferences generated between visually simple and visually complex pictures; that is, the less educated group generated more inferences for visually simple pictures than for visually complex pictures.

In summary, few studies have investigated inference generation broadly for graphic narratives. These studies reviewed in this section focus on developmental aspects of narrative comprehension with participant samples comprised of specific age groups. The focus of these studies is not on how graphic narratives as a modality affect reader comprehension processing, per se, but rather how the participants in the developmental stages they investigated generated inferences. The researchers used graphic narratives either because the materials matched the age group (e.g., Tompkins et al., 2012; Poulsen et al., 1979) or because, based on the Visual Ease Assumption discussed above, the researchers assumed non-linguistic graphic materials were a non-linguistic method of assessing inference ability (e.g., Ribeiro et al., 2010). Further researcher could investigate how visual modalities specifically influences what types of inferences and to what extent readers generate.

The results of the studies by Tompkins et al., (2012) and Poulsen et al. (1979) do support that narrative structure supports and facilitates readers' inferences about the narrative (Cohn, 2020). In the next section, researchers investigate how elements of narrative structure prompt readers to generate predictions of future events.

Predictions to Connect Current Events to Future Events

As discussed, anaphoric and bridging inferences tend to connect back to previous information. In this way, anaphoric and bridging inferences can be referred to as backward-oriented inferences because they connect back to prior information in the narrative (e.g., Magliano et al.,

1996). Inferences can also be future-oriented, such as in the case of predictions. With predictions, rather than connect to prior information, the reader predicts future information. For example, again, referring to the previously presented panel in Figure 2.1, a reader may think, “The man is going to sit down and order something to eat.” The reader generates this prediction based on their activated background knowledge of what people tend to do in diners. Readers can also generate predictions based on their understanding of narrative structure. For example, the reader may think, “This man must be the main character and right now this panel is just setting the scene. Soon some conflict will happen in this diner.” The reader generates predictions based on their recognition of narrative elements and their expectations for how a narrative tends to unfold.

Two studies investigating predictions during graphic narratives met the inclusion criteria for the current literature search. First, Cohn and Paczynski (2013) conducted five experiments on the role of agent (i.e., the one performing the action) or patient (i.e., the one receiving the action) characters in the generation of predictions. For example, if a man hands a plate of lasagna to his cat, the man is the agent performing the action of handing the plate, and the cat is the patient by receiving the results of the action. All participants were self-identified experienced comic readers. In their first experiment, adult participants viewed individual non-text panels in random order (i.e., panels were not intended to compose a narrative). Cohn and Pacyznski designated each panel as having either an agent or patient character and asked the participants to describe each panel with a single sentence. They then coded the participants’ sentence as active or passive voice. Participants tended to describe panels with agents in active voice and panels with patients in passive voice. Their second experiment was similar to their first experiment, but this time participants were asked to describe what would happen next for each panel in a single sentence.

They coded the participants' sentences as either predictable (i.e., more than one participant included the predicted event for that panel) or non-predictable (i.e., only one participant included that predicted event for a panel or responded with a non-prediction statement). Viewing panels with agent characters prompted participants to generate more predictions than panels with patient characters.

In their third experiment, participants saw six-panel comic strips without text one panel at a time at their own pace. In each narrative, Cohn and Paczynski identified critical panels in which the panel introduced some sort of relationship with an object (e.g., a boy with a ball), with a character in a passive state (e.g., sitting) or what they termed an unbounded action state (e.g., walking). These critical panels were edited into two separate panels (e.g., a panel with a boy and a ball separated into two panels, one panel showing the boy, one showing the ball). Participants then saw these narratives with the critical panel shown in original order (e.g., boy and then ball) or reversed order (e.g., ball and then boy). Cohn and Paczynski then measured participants how long participants viewed a single panel on a computer screen before moving on to the next panel. They found that the order of presentation of a character and an object did not affect participants' viewing times.

In their fourth experiment, Cohn and Paczynski used the same 30 six-panel comic strips, but this time edited the sequences so that either agents or patients were presented first. Procedures were similar to their third experiment, but unlike when they compared viewing times for characters and objects, this time they found that participants' viewing times were longer for agent panels than patient panels in general and this difference was greater when agents were presented first in the sequence. In their fifth and last experiment, Cohn and Paczynski replaced panels showing both an agent and a patient with either just an agent or just a patient. Critical panels

immediately preceding these panels then showed both characters with an action being performed. They found longer viewing times for critical panels after participants saw the patient panels; that is, participants processed actions faster after viewing the agent character. Overall, Cohn and Paczynski conclude that characters that fill a narrative role as agents in an event facilitate processing of the event and the generation of predictions. These findings support that readers' understanding of narrative structure facilitates their generation of predictions. In the next study, Cohn and colleagues (2017) extend these findings by measuring neural activity instead of participants' viewing times.

Cohn et al. (2017) manipulated initial panels in comic strips to include either a preparatory agent (i.e., action-doing character posed to do an action), a non-preparatory agent (i.e., action-doing character not posed for action), or patient. They measured neural activity though EEG for the initial panels as well as the peak panels following the initial panel. For this experiment, adults participants viewed comic strips one panel at a time. Cohn and colleagues observed greater frontal ERPs for preparatory agents than non-preparatory agents and patients during viewing of both initial and peak panels. That is, neural activity suggested that participants anticipated future actions or generated predictions of subsequent actions when viewing action-ready characters. The researchers concluded that in graphic narratives, agents may be "event builders" and help readers generate a prediction for the upcoming action or event in their mental model based on the preparatory agents.

At this point in the literature, studies that have investigated predictions in non-linguistic graphic narratives have focused on characters and their narrative roles guiding viewers to generate predictions. It is less known what other context and narrative cues prompt viewers of non-

linguistic graphic narratives to generate predictions or whether these cues differ from other modalities. Additionally, the reader may need to update their mental representation once they read information that confirms or contradict their predictions.

Updates to Correct or Confirm the Reader's Understanding

As the reader progresses through sequential images of the narrative, the reader takes in new information, and activates their background knowledge (Kintsch 1988; 1998; Magliano et al., 2013). In doing so, they generate a variety of online processes (i.e., various types of inferences) in order to construct a coherent mental representation of the narrative, specifically the situation model (Graesser et al., 1994; Zwaan et al., 1995). While constructing the situation model, the reader may need to update their understanding of the narrative because new information contradicts or confirms prior understanding (Cohn, 2020). For example, previously I discussed that readers may misinterpret visual information when assigning meaning to images. Instead of recognizing the setting of the panel in Figure 2.1 as a diner, the reader may think the man is on a bus. In later panels, the reader may recognize a table, a piece of pie, and a salt and pepper shaker set. At this point, the reader may think, "Oh, he's not on a bus. He's in a diner." Thus, the reader updated their situation model with an accurate understanding of the setting. As another example, after predicting that the man will order food, the reader may see a panel with a piece of pie and think, "Ok, he did order food." The reader updates their situation model to confirm their prediction. Currently no research has investigated the process of readers updating the types of processes they generate to develop a coherent mental representation while reading non-linguistic graphic narratives; thus, this is an empirical area that needs further investigation.

Summary. The studies included in this review which investigated online processing of graphic narratives support that readers of graphic narratives access visual information and generate anaphoric, bridging, emotion and internal state inferences, and predictions while reading (Cohn, 2020; Kendeou et al., 2020). The visual information about narrative content accessed influences the inferences readers generate (Magliano et al., 2013), as does the readers' understanding of narrative structure (Cohn, 2020).

It is important to note the processes not included in this discussion. Studies were not found that investigated the generation of elaborations or the process of updating mental representations during graphic narrative reading (Cohn, 2020; Kendeou et al., 2020). As already noted, studies that investigate anaphoric inferences are scarce, followed by studies investigating emotion and internal state inferences. So far, most of the research on processes during graphic narrative reading have focused on bridging inferences. Because viewers of graphic narratives have to "fill in the gap" across panels, bridging inferences are an important piece to understanding graphic narrative comprehension, and the focus on investigating bridging inferences is understandable. Still, the opportunity to investigate other types of processing is open and needed.

Additionally, few studies directly investigated the effects of modality on readers' online processing by comparing processing of graphic narratives with a linguistic modality, such as text. Therefore, less understood is whether reading a story as a graphic narrative, that is, with static images, activates the same prior knowledge as reading the same story as a text narrative. Subsequently, do readers generate the same inferences and predictions for a story whether viewing static images or reading text? More research is needed to understand graphic narrative comprehension and how it compares to comprehensions of other modalities.

Developing a coherent representation of a graphic narrative is the ultimate goal of graphic narrative comprehension. Once the reader has stopped reading a graphic narrative, there would be memory of the narrative. As mentioned previously, similar to text memory, readers of graphic narratives develop different levels of their mental representation. They develop the surface structure level representation (i.e., memory of the explicit visual information; Kintsch, 1988). They develop the image base representation (i.e., explicit meaning of the information presented in the narrative). Finally, they develop the situation model representation (i.e., information inferred from the information presented in the panels with their background knowledge) (Kintsch, 1988; 1998; Magliano et al., 2013). These mental representations of the narrative post-reading are referred to as comprehension products (McNamara & Magliano, 2009) and will be the next area discussed for this review of the literature.

Comprehension Products from Reading Graphic Narratives

While online comprehension processes are the in-the-moment cognitive processes a reader generates to construct mental representations as they read a narrative, offline comprehension products are the mental representations the reader has of the narrative post-reading or once the narrative is complete (Graesser et al., 1994; McNamara & Magliano, 2009). Products are the reader's memory of the mental representations constructed. In the literature search for studies investigating mental representations of graphic narratives, I included the terms *surface structure, text base, and situation model,* based on the mental representations identified in the CI Model (Kintsch, 1988; 1998; Kintsch & Van Dijk, 1978). None of the studies with a focus on comprehension products used these terms to frame their findings. Therefore, I organize relevant studies by the methods used to assess comprehension products, namely recall and recognition, and discuss what their findings suggest about offline graphic narrative comprehension.

Recall of Graphic Narratives

Recall involves participants retrieving information from memory to retell a story or saying (or writing) everything they remember from a story either read, read to them, or seen. In this way, researchers can ascertain the participants' memory based on what the participant includes in their recall. Few studies have used recall of graphic narratives, and only one, Poulsen et al. (1979), used graphic narratives without text.

In Paulsen and colleagues' (1979) study, after they asked 4-year-old and 6-year-old children to think-aloud while reading pictures stories without text, they asked the children to recall the stories just read. Across both age groups, the more children described the information from the pictures as a story with a narrative structure during the think-aloud portion, the more they included story information during the recall portion. Information described but not integrated into a story schema or narrative structure tended to not be recalled. This study supports that readers use narrative structure to integrate visual information into their mental representation of the story. For children who may not have as developed knowledge or understanding of narrative structure, it may be difficult to create a coherent representation of the narrative. While this study did investigate readers' offline products of non-linguistic graphic narratives, the researchers did not investigate the types of information (e.g., actions, emotions, events, etc.) the children tended to include in their recall of the story. The next study provides more detail on the types of story information recalled for visual stories.

In the next study, visual and linguistic information were presented together for participants to recall. Meringoff (1980) compared children's memory for animated film and picture book versions of the same story. Both the animated film and picture books had the same narra-

tion, but the film had more visual content in that the animated film had several minutes of dynamic visual presentation and the picture book included 17 static images. Participants included groups 6- to 8-year-olds and 9- to -10-year-old children. After viewing either the animated film or the picture books, Meringoff measured the children on post-viewing recall, picture ordering (i.e., putting in order 7 images of major events from the story), and asked comprehension questions that required inference generation post-reading. Children who saw the animated film scored higher on all inference questions. Children recalled more story vocabulary or figurative language from the picture book but recalled more actions from the animated film. Meringoff found no modality differences on picture ordering, but significant differences between age groups. Children who saw the animated film included more inferences based on visual information in their recall of the film, and children who read the picture book included more inferences based on textual information in their recall.

Meringoff acknowledged that the animated film had more visual content than the book version. Essentially, Meringoff tested children's linguistic narrative comprehension when paired with more or less and static or dynamic visual information. When more visual information is available, children recalled more actions but when less visual information is paired with the linguistic information, they tend to recall more linguistic information. These findings support that children did not have differences in their image base of the story (i.e., picture ordering) whether they saw the animated film or read the picture book. However, the children did have differences in their situation model representations depending on whether they saw the animated film or the picture book. The next set of studies use picture recognition, rather than recall, in order to investigate readers' image base and situation model representations.

Recognition of Visual Content of Graphic Narratives

Recognition tasks in graphic narrative comprehension studies ask participants post-reading to identify whether they had previously viewed an image within a graphic narrative. For example, as previously discussed, Magliano et al. (2017) used a recognition task to test participants' bridging inferences. Participants tended to misidentify images of bridging events that were missing from the narrative as having been previously seen. That is, when participants needed to generate a bridging inference to fill a coherence gap, they then falsely remembered seeing that missing image post-reading.

Additionally, Schmidt et al. (1979) investigated whether children integrate graphic narrative sequences and inferences into memory by testing their recognition of images from a visual sequence. Children participating in this study were either in kindergarten, first grade, or second grade. They saw five sets of 10 pictures which included six-panel narrative sequences and four distractor pictures. Children were told to look at the pictures and to try to understand the pictures and remember them. Then, children participated in a recognition task in which they answered two questions. First, children decided between two pictures at a time which one fit best with the story. Every pair either included an old picture from the story and a distractor image or a new picture of an action possibly inferred from the story and a distractor image. These new pictures all depicted actions that could be inferred from the story, but they varied in inferential distance. In a graphic narrative, inferential distance can be quantified by determining how many images or pictures would be needed to depict the inferred action from the originally depicted action. If an inferred action could be depicted in an image right after the original image, then it has less inferential distance than an action requiring three or four images from the original image. Then, for every pair of images, children also identified whether they had seen the picture before. Overall, children, especially older children, tended to not mistake new pictures as previously presented

pictures, especially for pictures with greater inferential distance from the original. This suggests that although children can and do generate inferences while reading graphic narratives, after reading, they are also able to recognize actions and events explicitly depicted in a narrative separate from images that fit or can be inferred. Thus, children are able to separate information included in their image base representation (i.e., memory for meaning and content explicitly depicted) and their situation model representation (i.e., memory for the content and generated inferences).

At first Schmidt et al. (1979) seems to contract Magliano et al. (2017) which found support for the memory distortion hypothesis, that the generation of bridging inferences leads to inaccuracy in memory for what was explicitly presented in the graphic narrative. However, perhaps the memory distortion hypothesis applies when the inferences in question are inferentially close to the explicit visual information shown in the narrative during reading and less applicable when the inferences are distant, as Schmidt and colleagues' findings suggest. Further research may test whether inferential distance should be taken into account for the memory distortion hypothesis.

Summary. The research on comprehension products of graphic narratives is much sparser than the research on comprehension processes. Aside from Magliano et al.'s (2017) study which focused on readers generation of bridging inferences during graphic narratives but incorporated a recognition task after reading, the studies included in this section on comprehension products are fairly dated. Additionally, Poulsen et al. (1979), Meringoff (1980), and Schmidt et al. (1979) all primarily focus on investigating developmental differences among younger readers. Currently there is little research investigating comprehension products of adult readers' graphic narrative comprehension. Based on the current search, little or no studies explicitly investigate

how non-linguistic graphic narrative readers develop a mental representation, especially when applying the levels of representation developed by the CI Model (Kintch, 1988; 1998). More research is needed to investigate these levels of representations in non-linguistic graphic readers and how these representations compare across modalities.

General Discussion from the Literature Review

The goal of this review of the literature was to find and discuss the previous research that has that investigated comprehension processes and products of graphic narratives identified by the CI Model (Kintsch, 1988; 1998; Kintsch & Van Dijk, 1978), the front-end/back-end framework (Magliano et al., 2013), the iLC Framework (Kendeou et al., 2020), and the PINS Model (Cohn, 2020). In doing so, I hoped to illustrate graphic narrative comprehension as a complex task and provide evidence that visual comprehension is not an inherently less complex than linguistic comprehension (e.g., Coderre, 2020). This search yielded 22 research studies that met the inclusion criteria for this literature search. The reviewed studies did not investigate all of the processes included in the search, such as elaborations and updates. Additionally, no study has investigated how visual narrative structure influences readers processing of graphic narratives. Similarly, no studies have specifically investigated the development of the mental representation of graphic narratives similar to what has been constructed by the CI Model (i.e., surface, image/text base, situation model). Thus, more research is needed to understand *how* readers construct their mental representations of graphic narratives and *what types* of mental representations are developed before additional conclusions can be formed about visual comprehension and how it may compare to linguistic comprehension.

What Processes Do Graphic Narrative Readers Generate While Reading?

The studies reviewed here support the theoretical frameworks of the CI Model (Kintsch, 1988; 1998; Kintsch & Van Dijk, 1978), the front-end/back-end framework (Magliano et al., 2013), the iLC Framework (Kendeou et al., 2020), and the PINS Model (Cohn, 2020). These frameworks and the research describe how readers process graphic narratives while reading. First, readers of graphic narratives take in visual information from the narrative through front-end processes (Magliano et al., 2013). The types of front-end processes readers engage in while reading depend on the modality, and for graphic narratives without any text, readers engage in scene gist processing, object/character processing, and sometimes motion processing. Collectively, these front-end processes and the action of taking in and interpreting visual information is called access (Cohn, 2020). Readers access the visual information in panels and assign meaning. How is a panel is framed or how much scene information is depicted can either help or hinder readers accessing the necessary information to comprehend the narrative (Cohn & Foulsham, 2020; Foulsham & Cohn, 2021). While readers may fixate their attention on important objects within the panel, contextual information in the scene is important for accessing the object. Additionally, information presented in a narrative order is easier for readers to access than information presented in random order (Foulsham et al., 2016). In addition to accessing visual information depicted explicitly, readers also interpret visual symbols that represent actions or information not explicitly depicted (Cohn & Maher, 2015).

Once readers take in visual information and assign meaning through front-end processing, they generate back-end processes to construct deeper meaning beyond the explicit visual information shown (Magliano et al., 2013). These back-end process include the generation of inferences to connect ideas across panels and expand upon visual information seen. Types of inferences include anaphoric, bridging, elaboration, and emotion (Kendeou et al., 2020). While only

one study has investigated anaphoric inferences for graphic narrative reading, it appears readers do recognize characters and objects seen previously. However, research has not yet investigated the extent to which readers identify previously seen objects and characters.

Bridging inferences connect panels by filling in missing actions that must occur between panels. Readers generate bridging inferences when a necessary action is not depicted (e.g., Magliano et al., 2016). These inferences are so salient in the reader's mental representation that they tend to remember inferred actions as being explicitly depicted in the narrative (e.g., Magliano et al., 2017). When readers need to generate a bridging inference, they search the panel for information to support the bridging inference (Hutson et al., 2018). Generating a bridging inference taxes the reader's working memory and cognitive resources by creating longer processing times (Magliano et al., 2016). Switching between modalities or types of information additionally taxes a reader's cognitive resources and impedes the generation of bridging inferences (Huff et al., 2020).

Readers generate other types of inferences as well to expand on the information presented. It appears children generate inferences on a variety of categories related to the narrative including goals and setting or place (Tompkins et al., 2012). Older adults' inferences based on visual information seem to be influenced by the complexity of the image and the older adult's years of education.

Children also generate inferences about characters' emotions (Gnepp, 1983; Poulsen et al., 1979; Tompkins et al., 2012) as well as characters' internal states (Lewis et al., 1994; Poulsen et al., 1979). Child readers generate emotion inferences based on characters' facial expressions, but as children age and presumably develop their socioemotional knowledge, they also

incorporate information from the situation and context to guide the emotion inferences they generate (Gnepp, 1983). Preschoolers can generate inferences about characters' internal states with guidance from an adult reader. Based on the current research, it is unknown the extent to which adult readers generate emotion and internal state inferences for graphic narratives.

Based on visual information accessed and expectations for how narratives tend to unfold (i.e., structure), graphic narrative readers also generate predictions (Cohn, 2020). So far, research supports that adult readers base their predictions on characters identified as those posed to complete an action (Cohn et al., 2017; Cohn & Paczynski, 2013). However, other narrative and contextual cues that may guide the generation of predictions have not been investigated with graphic narratives.

Lastly, as readers progress through the narrative, continuing to access new information as they read, graphic narrative readers may need to update their understanding of the narrative (Cohn, 2020). Perhaps a previous image was incorrectly identified, or a prediction was confirmed. A reader should update their mental representation with new information, but so far no study as investigated the extent to which readers update their mental representations. Other studies have mentioned readers' neural activity indicating updating in the context of studying how well readers access information (e.g., Cohn & Foulsham, 2020; Foulsham & Cohn, 2021; Foulsham et al., 2016).

What Products Do Graphic Narrative Readers Remember After Reading?

Readers generate cognitive processes while reading to construct their levels of mental representation (Kintsch, 1988; 1998; Magliano et al., 2013). Mental representations that readers have post-reading are referred to as comprehension products. It is posited that readers of graphic

narratives would construct surface, image base, and situation model levels of a mental representation. However, the research on readers' offline products of graphic narratives is sparse, and findings are limited to the amount of information recalled rather than the types of representations developed. Children tend to recall more information from a graphic narrative if while reading they organize visual information into a narrative structure rather than consider each panel as an individual image (Poulsen et al., 1979). Children also tend to recall more actions from visual narratives than they do narrative with more linguistic information. As discussed, the generation of bridging inferences may distort memory for explicit information depicted (i.e., the image base). That is, adult readers tend to remember inferred information integrated into their situation model representations as information explicitly depicted. Evidence with children suggests that this memory distortion is more likely to occur when the inferences are constrained by the narrative (Schmidt et al., 1979). That is, when an inference of an event is tightly guided by the narrative, the reader tends to remember that inference as explicit information. The further away from the narrative, the more a reader remembers the inferences as not explicitly stated.

There is much unknown about readers' offline products for graphic narratives. For example, it is unknown what type of mental representations are developed (i.e., surface, image base situation model; Kintsch 1988; 1998). Much research has been conducted that focused on the offline products of text narrative reading (e.g., Yeari et al., 2017; Hellmann, 2018); however, there are no studies to date that have compared the levels of the CI Model from recall across text and graphic narrative reading.

Limitations of the Literature Review

This literature review has two main limitations. First, I specifically searched and included studies that focused on graphic narrative comprehension; that is, comprehension of static sequential images such as comics and picture stories. Comprehension of other forms of visual narratives, such as film and animation, have been investigated (e.g., Kendeou et al., 2008; Magliano et al., 1996; Zacks et al., 2009). To have included these other forms of visual narrative would have provided a richer set of studies to review for visual narrative comprehension in general. However, my purpose was to understand the comprehension of visual content separate from linguistic or auditory content. Film and animation typically presents linguistic content through auditory dialogue. Even with the goal of searching for and finding studies that investigate visual comprehension separate from linguistic comprehension, some of the studies reviewed here used graphic narrative with text or auditory components (e.g., Lewis et al., 1994; Meringoff, 1980). Thus, this search yielded a much narrower lenses to understand visual narrative comprehension.

Second and relatedly, since I did limit the parameters of my review, studies are fairly limited; therefore, the conclusions from this literature review are also limited. Many of the studies discussed above provide evidence that readers of graphic narratives do generate many of the same processes that readers generate while reading text. Without direct comparison to text processing, it is only accurate to conclude that readers generate similar processes across modalities, not how modality affects processing or whether readers generate processes to the same extent across linguistic and visual narrative modalities. Thus, investigating the processes and products of both text and graphic narrative modalities in the same study is warranted.

Finally, this review highlights how little research has investigated offline comprehension products of graphic narratives. However, the studies reviewed here include only children partici-

pants and also do not explicitly test readers' levels of mental representation (i.e., image base, situation model). Therefore, it is difficult to generate conclusions about readers offline comprehension products for graphic narratives. Identifying these limits provides a guide for implications for additional research.

Implications for Research

In light of the limitations of the review, the implications for research are two-fold. First, there are many processes studied with text reading that have not been studied with graphic narratives (e.g., global and local inferences; van den Broek et al., 2001). Of the terms searched for this review, elaborations and updates returned no results. Similarly, the text base and situation model representations identified in the CI Model (Kintsch, 1988; 1998) have been studied for text comprehension (e.g., Wannagat et al., 2017) but understudied for graphic narrative comprehension. Therefore, the research is needed to thoroughly compare how readers process text and graphic narratives and how they remember and comprehend text and graphic narratives post-reading.

Second, some research not discussed in this review has directly compared visual comprehension with linguistic comprehension with neurodiverse populations (e.g., Coderre, 2020; Coderre, 2018). These studies show a lack of support for the Visual Ease Assumption, but this assumption can also be tested with other populations. Directly comparing processing and products for text narratives and non-linguistic graphic narratives would inform the field how visual comprehension compares to linguistic comprehension and whether the form of information (i.e., text, pictures) influences how people think about and comprehend narrative information.

Implications for Practice

The iLC Framework (Kendeou et al., 2020) is designed to facilitate interventions to scaffold children's inference skills with text narratives through visual narratives. Studies support that

readers have a general inferencing ability that spans across modalities (e.g., Kendeou et al., 2008); thus, if a reader's inference skills are developed through graphic narrative reading, that skill should transfer to text reading. While many of the studies here do not compare inferences generated during graphic narrative reading and text narrative reading, they do support that readers generate similar processes for graphic narratives that readers generate for text. The previous studies that support a general inference ability compared narrative modalities all of which included linguistic information. Thus, it is difficult to conclude whether readers have a general inference ability or a linguistic inference ability and whether this distinction has practical significance for intervention. Without further research, it is unknown whether modality influences how a reader exhibits their general inference ability. That is, does a proficient comprehender generate inferences for both text and graphic narratives? Or does modality prompt readers to generate similar types of inferences to different extents? If so, how then does scaffolded inferencing skills transfer across modalities? Further research directly across linguistic and visual narrative modalities would address these questions and inform the iLC Framework and related interventions.

Based on the limitations in the literature highlighted by this review and the need for both research and practice to directly compare comprehension processes and products of linguistic and visual modalities, I conducted the following study.

Overview of the Study

In this study, I examined and compared both comprehension processes while reading (i.e., online) and comprehension products post-reading (i.e., offline) across text and non-linguistic graphic narratives. Guided by the CI Model (Kintsch, 1988; 1998; Kintsch & Van Dijk, 1978), Magliano et al.'s (2013) front-end/back-end framework, the PINS Model (Cohn, 2020),

and the iLC framework (Kendeou et a., 2020), I conducted a study involving think-aloud protocols to investigate comprehension processes and recall tasks to investigate comprehension products. Participants read and recalled both text and non-linguistic graphic narratives (i.e., the text narratives were written based on their non-linguistic graphic narrative counterparts; See *Materials* section for additional information). Text narratives and non-linguistic graphic narratives were chosen in order to investigate the effects of linguistic and visual information (i.e., modality), respectively, on readers' narrative comprehension. Having participants read both text and non-linguistic graphic narratives extends previous research by manipulating the front-end processes participants engage in (i.e., processes involved in decoding and taking in information) for each narrative condition and subsequently test the effects of modality differences on the various processes generated during reading (i.e., back-end processes) and on the products or levels of mental representations developed after reading (i.e., text base and situation model). I also controlled participants' text print exposure (Acheson et al., 2008; Stanovich & West, 1989) and visual language fluency (Cohn, 2014) as covariates in my analyses because I am interested in the effect of modality on comprehension processes and products, not variability in participants' familiarity with text and graphic narratives.

Study Purpose

The purpose of the current study was to incorporate each of the above theories as a foundation for investigating whether readers generate comprehension processes for matching text and non-linguistic graphic narratives to the same extent and if readers produce similar or different offline products for text and non-linguistic graphic narratives when controlling for reader text print exposure and visual language fluency. This study examined the effects of modality on reader processing and products of comprehension, not prior experience nor familiarity with a modality.

That is, the aim of this study was to examine and identify differences in narrative comprehension (processes and products) due to modality.

Assumptions and Limitations

For this study, I assume that first, both reading text and viewing sequential visual images (i.e., graphic narratives) involve decoding and ascribing meaning to information decoded; that is, while I am interested in differences in comprehension processes and products across modalities, I assume that both text and graphic narratives involve essentially the same act of reading. Second, for this reason, I assume that one modality is not necessarily "easier" to comprehend than the other. Thus, this aspect is in opposition to the Visual Ease Assumption (e.g., Coderre, 2020). The purpose of this study is to explore whether differences across modality lead to differences in processes and products, not that differences in modality leads to "better" comprehension for one modality over the other.

I also assume that the text and non-linguistic graphic versions of the narratives used in this study communicate the same information. That is, the materials used for this study were developed by Marvel for publication and included the same story line. Thus, because these materials were not developed by myself, I can only assume the original intent of the authors was to ensure that the text and graphic versions included the same information. However, this assumption consequently leads to a limitation of this study. I recognize the limits of creating absolute equivalent text and visual versions of the same story. Colloquially speaking, a sentence may contain about 15 words, while "a picture is worth a thousand words." That is, a visual image include and communication more information than a single sentence of text. For example, a reader may see in a panel of the graphic version of a story that a character has two matching lamps in her bedroom, but that information is not pertinent to the character's goals and events of a story and is

therefore not included in the corresponding sentence in the text version. Thus, it should be noted that additional information could have been included in the text versions; however, the authors may have thought not all information in an image was necessary to understand the story and, for this reason, may have been deliberately left out of the text version. This could potentially create a negative impact to the reader's comprehension.

In that same vein, pertinent information may be presented in a visual image, but a reader may fail to notice that information. For example, an image may include an object integral to the story, but the reader may attend to other, non-pertinent information in the image rather than attend to the integral object. In the corresponding text sentence, the object and its relevance to the story could be explicitly stated. Therefore, relevant information could be explicitly stated in the text, but for an image, a reader may or may not attend to similar relevant information. For these reasons, there are limitations to creating perfectly matched versions of a narrative with text and visual modalities and still present narratives readers may encounter outside the lab. Acknowledging these limitations, the *Narratives* section in the Methods chapter (Chapter 3) provides detailed explanation for how the text and non-linguistic graphic versions of these narratives were obtained as well as how they are matched.

Additionally, this study involves a small sample size, although typical for the type of study and the methods used (i.e., think-aloud and recall; see the *Power Analysis* section in Chapter 3). This small sample consists of college students from one university in the southeast, all of whom are relatively skilled readers. Future directions in the Discussion chapter (Chapter 5) discusses how these methods could be extended to younger readers, readers of varying comprehension profiles, and readers with language and/or emotional/social impairments (e.g., individuals on the Autism Spectrum Disorder). The purpose of this study was to explore the extent to which

readers may generate comprehension processes and products across modalities, and I acknowledge that is a broad purpose. However, no one study can tell the full story of a construct or behavior. Thus, I conceptualize this study as a starting point for a specialized area of research.

Significance of the Study

To my knowledge, no study has investigated the types of comprehension processes generated during reading and products produced after reading across *text* and *non-linguistic graphic narratives*. In addition, there has been no study that has examined the specific types of processes during reading across both text and non-linguistic graphic narratives (Kendeou et al., 2020; Coderre, 2020; Cohn & Magliano, 2020), and the types of products produced after reading (Kintsch, 1988; 1998; Magliano et al., 2018). Finally, because text print exposure and visual language fluency have both been shown to predict successful reading skill of their respective modalities (e.g., Acheson et al., 2008; Cohn & Maher, 2015), the current study controlled for reading skill for each modality in order to directly examine whether comprehension processes and products differ across modalities.

The results of this study answer the call to investigate the similarities and differences in the extent to which readers generate online processing and offline products across narrative modalities, specifically, text and non-linguistic graphic narratives (e.g., Cohn, 2020; Kendeou et al., 2020; Cohn & Magliano, 2020). By comparing matched text and non-linguistic graphic versions of narratives and controlling for familiarity of both modalities, any differences in processing and products should be attributed to modality. This study was designed not to conclude whether comprehension for text and graphic narratives is broadly similar or different, but to investigate the nuances of similarities and potential differences across the two modalities. For example, if results indicate a significant difference in either think-aloud or recall responses (or both) across the

modalities, the results of post hoc analyses focus on *in what way* the think-aloud or recall responses differ across the modalities.

Whether results of this study show significant statistical differences in narrative comprehension or do not, the methods used to understand readers' cognitive processing and comprehension across narrative modalities will be expanded. While a few studies have used think-aloud tasks with non-linguistic graphic narratives (e.g., Magliano et al., 2016; Magliano et al., 2017), those studies looked at participants' responses at specific points in the narrative, not throughout the whole narrative. Thus, this study examines how readers process and construct their representations of text and non-linguistic graphic narratives throughout the narrative. Additionally, no studies have used recall tasks for non-linguistic narratives. The current study is the first to study readers' image base and situation model representations through recall of completely visual narratives. Results of this study lay the foundation for further studies investigating narrative comprehension across modalities.

Therefore, the following research questions are:

Research Questions

1) To what extent do readers differ in the generation of different types of comprehension processes while reading text and non-linguistic graphic narratives after controlling for text print exposure and visual language fluency?
2) To what extent do readers differ across comprehension products for text and non-linguistic graphic narratives after controlling for text print exposure and visual language fluency?

Hypotheses

The hypotheses addressing each research question are:

Research Question 1. To what extent do readers differ in the generation of comprehension processes while reading text and non-linguistic graphic narratives after controlling for text print exposure and visual language fluency?

H0: There are no mean differences in the type of online comprehension processes readers generate as observed in think-aloud tasks across text and non-linguistic graphic narratives after controlling for participants' text print exposure and visual language fluency.

H1: Main: There are mean differences in the type of online comprehension processes readers generate as observed in think-aloud tasks across text and non-linguistic graphic narratives after controlling for participants' text print exposure and visual language fluency.

The expectations for this hypothesis are that readers of both text and graphic modalities should generate anaphoric, bridging, elaborative, local/global, predictive, emotion, updating, and internal state inferences (Cohn, 2020; Kendeou et al., 2020). Based on previous findings, participants should generate more anaphoric, bridging, elaborative, emotion, global/local, predictive, and updating inferences during reading of non-linguistic graphic narratives than text narratives (e.g., Coderre, 2020; Cohn & Magliano, 2020; Gnepp, 1983; Loughlin et al., 2015). On the other hand, participants should generate more internal state inferences during reading of text narratives than non-linguistic graphic narratives (e.g., Coderre, 2020; Cohn & Magliano et al., 2020; Magliano et al., 2013; Magliano et al., 2018). It is also expected that participants should generate more access statements during reading of non-linguistic graphic narratives than text narratives (Cohn, 2020).

It is expected that participants should also exhibit evidence of considering narrative structure as they think-aloud for both narrative modalities (Cohn, 2020). Participants should generate more goal statements during reading of text narratives than non-linguistic graphic narratives

(e.g., Johnson & Mandler; 1980; Rumelhart, 1977; Stein & Glenn,1979). Participants should also generate more event and scene statements for non-linguistic graphic narratives than text narratives (Cohn, 2013; Cutting, 2016; Magliano et al., 2013). These expectations will be conducted through post-hoc analyses for mean differences found for the main hypothesis.

Research Question 2. To what extent do readers differ across comprehension products (i.e., text/image base and situation model) for text and non-linguistic graphic narratives after controlling for text print exposure and visual language fluency?

H0: There is no difference across reader comprehension products as measured by participants' recall across text and non-linguistic graphic narratives after controlling for participants' text print exposure and visual language fluency.

H2: Main: There are differences across reader comprehension products as measured by participants' recall across text and non-linguistic graphic narratives after controlling for participants' text print exposure and visual language fluency.

The expectations for this hypothesis are that readers of both text and graphic modalities should generate comprehension products that reflect the text/image base and situation model levels of representations (Kintsh, 1988; 1998). Additionally, regarding participants' text/image base representations of the narrative, participants should include more accurate story information and more accurate description statements in recall of non-linguistic graphic narratives than text narratives (Loughlin et al., 2015; Magliano et al., 2013). Regarding participants' situation model, participants should include more valid action inferences, valid emotion inferences, and valid general inferences in recall of non-linguistic graphic narratives than text narratives (Cohn, 2013; Cutting, 2016; Magliano et al., 2013). On the other hand, participants should include more valid goal in-

ferences and valid internal state inferences in recall of text narratives than in recall of non-linguistic graphic narratives (Johnson & Mandler; 1980; Magliano et al., 2013; Rumelhart, 1977; Stein & Glenn,1979). The findings for these expectations are reported below.

CHAPTER III

METHOD

Conceptual Framework

Research Design

The goal of this study was to compare the extent to which readers generate online processes and offline products across text and non-linguistic graphic narrative reading. Therefore, for this study, the independent variable was narrative modality: text and non-linguistic graphic narratives. Because all participants saw both graphic and text narratives, this study was a within or repeated measures design. Participants' Author Recognition Test (ART; Acheson et al., 2008) and Visual Language Frequency Index (VLFI; Cohn, 2014) scores were also included as covariates to control for participants' text print exposure and visual language fluency. Therefore, any differences in outcome measures would be due to modality. The study had two kinds of outcome measures or dependent variables: (1) online process dependent variables as measured by coding think-aloud responses stated during reading, and (2) offline product dependent variables as measured by coding participants' recall of the narratives post-reading.

The order and modality of narratives was counterbalanced across participants. Every participant saw all four narratives (listed as Narratives A, B, C, and D in Table 3.1 below; also see Appendix A with the texts listed in A-D order) in different orders. Every participant also saw two of the narratives presented in text form and two of the narratives presented in non-linguistic graphic form. Participants were randomly assigned whether they saw the text or graphic modality condition of each narrative first. See Table 3.1 for an example of the random assignment and counterbalance order of the narratives for a portion of the participants.

Table 3.1.

Example of Random Assignment and Counterbalance of Narrative Presentation by Narrative (Narr) and Modality (Mod)

Participant	Narr	Mod	Narr	Mod	Narr	Mod	Narr	Mod
1	A	text	B	graphic	C	text	D	graphic
2	B	graphic	C	text	D	graphic	A	text
3	C	text	B	graphic	A	text	D	graphic
4	D	graphic	C	text	B	graphic	A	text
5	A	graphic	B	text	C	graphic	D	text
6	B	text	A	graphic	D	text	C	graphic
7	C	graphic	B	text	A	graphic	D	text
8	D	text	B	graphic	C	text	A	graphic
9	A	text	D	graphic	C	text	B	graphic
10	B	graphic	A	text	D	graphic	C	text
11	C	text	A	graphic	B	text	D	graphic
12	D	graphic	B	text	C	graphic	A	text
13	A	graphic	D	text	C	graphic	B	text
14	B	text	D	graphic	A	text	C	graphic
15	C	graphic	A	text	B	graphic	D	text

Power

Power was estimated using an priori power analysis to determine the optimal sample for this study. For this power analysis, a multivariate analysis of covariance (MANCOVA; see Data Analysis in Chapter 4: Results) was used. However, a power analysis using a MANCOVA is a complicated statistical procedure. Specifically, a power analysis for a MANCOVA would need to consider the matrix of relationships among predictors, among all dependent variables, of predictors to dependent variables, and of covariates to predictors and dependent variables. Therefore, based on recommendations found when searching for other power analysis methods (based on the literature and personal communication with committee member, Dr. Hongli Li), I estimated an appropriate sample size by conducting power analyses for a variety of tests related to a MANCOVA to gauge a range of possible sample sizes (see Table 3.2 below). All power analyses

were, thus, calculated using G*Power 3.1 (Faul et al., 2009). First, I calculated the estimated sample size for an analysis of covariance (ANOVA) with two groups (i.e., text and graphic) with a medium effect size of .25 (Cohen, 1988; 1992). Based on these analyses, a calculation was estimated for a sample size of $N = 210$ needed for this study. Second, I calculated two estimated sample sizes for multivariate analysis of variance (MANOVA) for within factors for two groups (i.e., text and graphic) with a medium effect size of .25 (Cohen, 1988; 1992). The first MANOVA in this calculation included 16 dependent variables (i.e., number of codes for the think-aloud task; see Table 3.7 below) and the second included 14 dependent variables (i.e., number of codes for the recall task; see Table 3.8 below). Each calculation estimated the same sample size of $N = 44$ needed for this study. Thus, given the differences between the estimated sample sizes from the two power analyses, a recommended range for my sample size was from $N = 44$ to 210. To determine which size would be most optimal, previous studies involving think-aloud and recall tasks were reviewed. Based on this review, narrative comprehension studies involving think-aloud and recall tasks often involve sample sizes ranging from approximately 20 to over 100 participants (e.g., Botas, 2017; Kraal et al., 2018; Magliano & Millis, 2003; Magliano et al., 2016; Magliano et al., 2017; Narvaez et al., 1999; Renz et al., 2003; Wright et al., 2011). Thus, based on the above power analyses, recommendations, and the review of this type of research, a final sample size range was chosen as $N = 50$ to 100 participants.

Table 3.2.

Estimated Sample Size Across Statistical Tests

Test	Groups	Dependent Variables	Effect Size	Estimated *N*
ANCOVA: Fixed effects, main effects and interactions	2	1	.25	210
MANOVA: Repeated measures, within factors	2	16	.25	44
MANOVA: Repeated measures, within factors	2	14	.25	44

Participants/Setting

This study was conducted at Georgia State University (GSU). Specifically, undergraduate student participants were recruited through GSU's Department of Learning Science's (DLS) online research scheduling system (i.e., SONA). SONA is a voluntary sign-up system for student participation in research studies through the DLS. Participants received credit required for their coursework after completing studies for which they signed up for. The amount of credit awarded for participation in a study depends on the time spent participating. The current study required an hour of a participant's time, and therefore, participants were awarded one research credit for their course. Participants signed up for a day and time to participate in the study and then received a link to participate virtually in a Zoom meeting. Sixty-three participants signed up for this study, and 51 participants completed the study. Participants who did not complete the study either decided to end the session early or skipped tasks, or an error occurred with the recording resulting in missing data. Of the 51 participants, three participants skipped the recall task for at least one narrative, and their data was excluded from the recall analyses. Therefore, think-aloud analyses were based on a sample of $N = 51$ participants, and recall analyses were based on a sample of $N = 49$ participants.

A voluntary demographic survey in Qualtrics was presented at the end of the virtual session in order to collect participants' age, ethnicity, gender identity, and class standing. Of the 51 participants who completed the study, 42 (66.7%) also completed the demographic survey. The mean age of participants who completed the demographic survey was 25.55 (*SD* = 9.55), with a range of 18 to 55 years old. The descriptive statistics of those participants who completed the survey are listed in Table 3.3 below. The survey also asked participants to rate their experience and expertise both reading and creating text narratives outside of their educational classes (see Table 3.4). These items were worded similarly to the items assessing visual narrative experience in the VLFI (see Measures and Materials below).

Table 3.3.

Demographic Information

Category	Frequency	Percentage
Gender Identity		
Man	7	16.7
Woman	33	78.6
Non-binary / gender non-conforming	2	4.8
Racial Identity		
Asian	4	9.5
Black	20	47.6
White	16	38.1
Other	1	2.4
Prefer not to answer	1	2.4
Ethnicity		
Hispanic	4	9.5
Non-Hispanic	38	90.5
Class Standing (*N* = 40)		
Freshman	1	2.4
Sophomore	6	14.3
Junior	9	21.4
Senior/Graduate	24	57.1

*Note: *N* = 42.

Table 3.4.

Participants' Text Experience

Variable	*M(SD)*
Current expertise reading text narratives	3.88(9.55)
While growing up	4.17(.82)
Current text narrative writing ability	3.57(.91)
While growing up	3.50(1.02)
Age when began reading text narratives	8.21(3.04)
Age when began writing text narratives	9.78(3.50)

Note: *N* = 42. Rating on a Likert scale of 1 to 5: 1 = below average, 2 = slightly below average, 3 = average, 4 = slightly above average, 5 = above average

Screening Measures and Procedures

The Letter Word Identification subtest of the Woodcock-Johnson III (WJ-III; Woodcock et al., 2001; 2014) Test of Achievement was used to measure participants' text decoding skills and ensure that participants were able to complete the study tasks. For this study, participants were required to read text narratives at the fourth to fifth grade level. Therefore, participants were first screened for their decoding skills to ensure text decoding ability would not influence variation in outcome measures across text and non-linguistic graphic narratives. For this subtest, examinees read aloud a list of increasingly difficult words until they no longer pronounced the words correctly or completed the list.

Participants in this study started with the suggested starting point for Grade 5 through Average Adult at item 46. To begin the administration, the researcher asked, "What is the word at the top of the list?" After the participant responded, the researcher said, "Please read the rest of the words." If the participant read the first six words incorrectly, the researcher proceeded in order to prior items until the participant read six words in a row correctly. Because this study required participants to read text narratives at the fifth grade reading level, the subtest was discontinued if a participant could not answer six items correctly past item 30, the suggested starting point for Grade 3. Of the 63 participants who signed up for this study, all met this benchmark. If the participant read the first six words correctly, then the researcher proceeded to the set of words. The researcher repeated the instructions to read the word at the top of the list and then the rest of the words until the participant answered six words in a row incorrectly or completed the final item.

Every correct response on the WJ-III Letter Word Identification subtest was scored with a 1. Words must be pronounced correctly and fluently (i.e., smoothly) to receive credit for the

word. Every incorrect response was scored with a 0. All items below the basal score were scored as correct. Raw scores were used in this study to screen and describe participants. Raw scores for this subtest range from 0 to 76. See Table 4.1 in the Results chapter for descriptive statistics for the Letter-Word Identification raw scores of recruited participants included in this sample.

Measures and Materials

Covariate Measures

The goal of this study was to test whether there are differences in narrative processing and comprehension products due to modality. Prior experience with a modality can influence a reader's processing and comprehension of a narrative in that modality. For example, more experienced comic book readers comprehend graphic narratives more easily than less experienced comic book readers (Cohn, 2021). I included two covariate measures in order to control for participants' background experience with text and graphic narratives. The Author Recognition Test is a measure of exposure to print, and the Visual Language Fluency Index is a measure of experience with graphic narratives.

Author Recognition Test. Scores from the Author Recognition Test (ART; Acheson et al., 2008; Stanovich & West, 1989) quantifies an individual's exposure to print, specifically narrative text. The ART includes a list of names on a single page, including names of literary and contemporary fictional authors. Participants identified which names are names of authors. Scores for the ART are obtained by counting a participant's total correctly identified authors and names incorrectly indicated as authors. Then, a final score is derived by subtracting the total incorrect authors identified from the total correct authors identified. See Table 4.1 in the Results chapter for descriptive statistics of this sample's ART scores.

Visual Language Fluency Index. Scores from the Visual Language Fluency index (VLFI; Cohn, 2014) quantifies an individual's experience with visual narratives, including comic books, comic strips, graphic narratives, film, and animation. Participants rate their experience and expertise both consuming and creating visual narratives both as a child and at their current age. *VLFI.* The index's creator Neil Cohn provides a formula for calculating a participants' over-all visual language fluency score based on their responses to individual items. The formula for deriving participants' overall visual language fluency score is as follows:

(Mean Comic Reading Frequency x Comic Reading Expertise) + [(Comic Drawing Fre-quency x Drawing Ability)/2]

See Table 4.1 in the Results chapter for descriptive statistics of this sample's VLFI scores.

Experimental Materials: Narratives

Graphic Versions. The four narratives for this study included selections from Marvel Comics' "'Nuff Said" comic book issues. In 2002, Marvel Comics writers and artists completed a challenge to create an issue of their series without incorporating any text. Therefore, the stories in these issues include visual-only (i.e., non-linguistic) graphic narratives. In addition to the chal-lenge in 2002, writers included text-written story plots for the comics and directions to the artists in the second half of the issue to illustrate how a comic book story evolves from the writer's written directions to a completed comic book issue. Most, if not all, of these text directions are incomplete within the issue. That is, the comic books issues only shared a portion of the text di-rections. Readers were directed in 2002 to read the complete text on Marvel.com. Per my knowledge, these complete text directions are no longer listed on Marvel.com. Thus, each issue included text for the first half of the story. In order to present both non-linguistic graphic and text modalities of each narrative, then, narratives for this study were limited to beginning events in

the issues because these texts were complete for the visual only versions. In addition, the issues were chosen based on whether readers with limited background knowledge about the characters and storylines could understand the story. For example, the *Avengers* issue was not chosen for this study because readers need to know events in prior issues to understand the events in the current issue. In addition, stories within other issues were deemed overly complicated for the purpose of this study. For example, in the *X-Men* issue, two characters telepathically enter another character's mind to explore his memories and consciousness. Expecting a naïve reader to infer that these characters are capable of these goals and actions was beyond the scope of this study. Despite instructions to not use texts in their issues, some creators still included some text or text-based symbols (e.g., a "?" in a character's thought-bubble or a written letter providing context), and then there were other issues which included excessive text (i.e., more panels with text than what could reasonably be edited to create a coherent narrative). These were also not chosen for this study.

Based on all these considerations, the narratives in this study included the text and the visual only versions and were: *Amazing Spider-Man* volume 2, issue 39 (Straczynski et al., 2002), *Captain America* volume 3, issue 50 (Jurgens et al., 2002), *The Incredible Hulk* volume 2, issue 35 (Jones et al., 2002), and *Thor* volume 2, issue 44 (Jurgens et al., 2002; see Appendix A). These issues were chosen because they were determined to not require readers to need background knowledge about the characters and event sequences to understand the events. Again, narratives for this study were limited to panels and events from the pages the text directions covered. Then, I also shortened the narratives by leaving out superfluous panels (e.g., multiple panels including different viewpoints of the same setting). If the sequence was coherent without a panel, that panel was edited out of the narrative. This editing was done to ensure that narratives

were of similar length and short enough for participants to read completely in the time allotted. The narratives were also abridged in a way so that the story ended with some sort of resolution to maintain a coherent narrative structure. Narratives were abridged to ensure that each had a clear beginning, middle, and end and had multiple event sequences.

As mentioned, panels with any language (e.g., description, dialogue, thought bubbles, hand-written letters) were not included; however, some panels do include words that appear in the visual images themselves (e.g., the word “Restroom” on a restroom door; a name on a desk nameplate). Any word that did appear in a panel was also included in the corresponding text version of that narrative.

Currently, a readability scale for non-linguistic graphic narratives does not exist; therefore, readability for each narrative was based on their text versions (see Text Versions below). The average length of the graphic narratives is 30 panels with the shortest at 29 panels and the longest at 31 panels. Narrative A (i.e., *Amazing Spider-Man*) was 29 panels long; Narrative B (i.e., *Captain America*) was 30 panels; Narrative C (i.e., *Incredible Hulk*) was 31 panels; and Narrative D (i.e., *Thor*) was 29 panels.

Text Versions. Once the panels for the graphic versions of each narrative were chosen, the text was then matched with the graphic panels to modify the text versions of each narrative. I included one sentence per panel and also retained as much original text from the writers as possible. Text directions such as “open shot, panning up the apartment” were not included. In addition, if the writers wrote in present tense, the text was changed to past tense. If the writers made explicit something that someone would have to infer in the graphic version, that information was also not included. While names of characters are not made explicit in the graphic versions, names were made explicit in the text versions.

Thus, for this study, I modified text versions of the narratives by editing the writers' directions provided in the issues or writing a sentence for each panel. The text versions of the narrative were modified to not include specific directions to the artists (e.g., "I want her looking sad and forlorn here") and to be of equivalent reading level across narratives. Thus, I presented both text and non-linguistic graphic versions of each narrative (see Figure 3.1 for an example).

The text and non-linguistic graphic versions of each narrative contained the same number of panels and sentences. That is, for every panel shown in the non-linguistic graphic version of a narrative, there was a corresponding sentence in the text version of the narrative (see Figure 3.1 for an example and Appendix A for the complete set of the materials). Readability for the text versions will be determined by Flesh-Kincaid (FK; Kincaid et al., 1975). The average readability level as measured by Flesh-Kincaid (FK; Kincaid et al., 1975) of the text versions of the narratives is 4.5, ranging from 3.9 to 5.3. Text readability of the text version of Narrative A was 5.3, Narrative B was 4.0, Narrative C was 4.4, and Narrative D was 4.4.

By using published comic books as materials for this study rather than narratives created for this study, the current study targeted the comprehension processes and products during narrative comprehension for naturalistic reading. Each of comic book issues for which the materials are from was published in 2002, approximately 20 years ago, and none of the issues are particularly popular among comic readers; thus, there are no recency concerns.

Figure 3.1.

Examples of Corresponding Non-linguistic Graphic Panels and Text Sentences

Slide 1

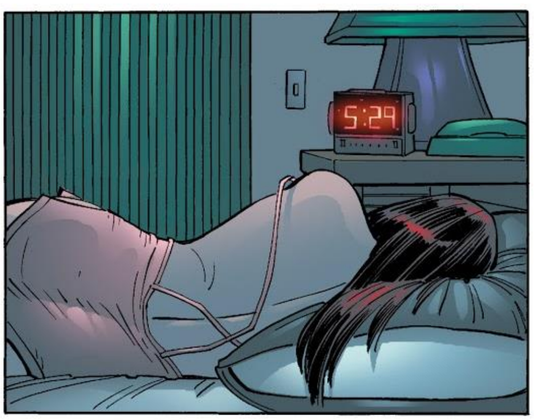
Graphic Version

Text Version

Mary Jane lay on one side of the bed alone.

Slide 2

Graphic Version



Text Version

The alarm clock read 5:29

.

Note: Images are from *Amazing Spider-Man* volume 2, issue 39 (Straczynski et al., 2002).

Participants saw either the non-linguistic graphic or text version of each of the four narratives, with each participant seeing two narratives as non-linguistic graphic and two narratives as text. Because some participants were excluded from the analyses, the remaining sample of participants saw some narratives more frequently as non-linguistic graphic and some narratives more frequently as text (see Table 3.5). A chi-square difference test ($\chi^2(3, 51) = .863$, $p = .843$) indicates no significant differences in how frequently participants saw a narrative as either text or non-linguistic graphic.

Table 3.5.

Distribution of Non-Linguistic Graphic and Text Versions Across the Narratives

	Modality	
Narrative	Non-Linguistic Graphic	Text
A	26	25
B	24	27
C	24	27
D	28	23

Note: N = 51.

Procedures

This study was conducted virtually. Participants signed up for the study and scheduled a virtual, synchronous session through the GSU DLS SONA system. Through the GSU DLS SONA website, participants received a Zoom link to enter their virtual research session. Participant responses were recorded throughout their virtual session vis Zoom, which was later downloaded for transcription and coding purposes. Participants were made aware of the recording at the beginning of the session and orally provided informed consent before recording began. Once they completed the study tasks in the virtual session, they received SONA research participation credit.

ART

After the participants completed the screening measure and remained eligible to participate in the study tasks, they first completed the ART. For the ART (Acheson et al., 2008), the researcher screenshared the list of names and provided instructions. The instructions for the test traditionally administered in-person were edited to suit the virtual format of this study. Researchers instructed participants, "Below is a list of names. Some of them are authors of books, and some of them are not. Please read the list silently to yourself and say out loud for the recording the ones you know for sure are authors. There is a penalty for guessing, so you should only say out loud those names about which you are absolutely certain. You can let me know when you are finished." The researcher then allowed the participants remote control of the screen in order to scroll through the list at their own speed. Participants silently read a list of names to themselves and orally indicated from a list of names which names they recognized as authors. Once the participant indicated they identified all the authors they recognized, the researcher continued to the next measure, the VLF index.

VLFI

For the VLF index (Cohn, 2014), the researcher screenshared the index and gave participants context that these questions would be about their experience with visual narratives. The researcher read aloud the questions in the index and explained the rating scales provided for each item. Participants answered orally. Once the participant completed the items in the index, the researcher continued to the narrative tasks.

Narrative Tasks

For the think-aloud and recall tasks, participants saw four narratives, two text and two non-linguistic graphic narratives (see Appendix A). Before seeing the first narrative, the researcher demonstrated the think-aloud and recall tasks with two practice narratives, one text and

one non-linguistic graphic narrative. For the first practice narrative (text), the researcher demonstrated thinking aloud and recall. For the second practice narrative (graphic), participants were prompted “ok, what would you think aloud for this one?” on the last slide and asked to recall the story. Then the researcher introduced the first narrative for that participant.

Think-Aloud. For the think-aloud task, participants saw narratives one panel or read one sentence at a time. For text narratives, they were prompted to read aloud the sentence and then say aloud whatever comes to mind. For non-linguistic graphic narratives, they were prompted to view the panel and then say aloud whatever comes to mind. Participants were told to say whatever they are thinking, even if it seems obvious, because there is no right or wrong answer. Participants were given remote control of the shared screen through Zoom and progressed through the narratives at their own pace by pressing a spacebar or arrow key. They were also instructed they could not go back to previous panels or sentences. Responses were audio recorded for transcription and coding. In between the think-aloud and recalls tasks for a narrative, participants saw and completed a distractor task (i.e., two simple math problems) to eliminate any potential recency effects from reading.

Recall Task. After thinking-aloud, participants were prompted to complete the distractor task (i.e., answer two simple math problems). After the distractor task, they were prompted to recall as much of the narrative they just viewed or read. Participants were directed to say all that they remember out loud. After each narrative’s recall task, participants answered two comprehension questions written for each narrative. Recall responses will be audio recorded for later transcription and coding.

After completing think-aloud and recalls for all narratives read, participants were asked whether they had read the comics before and to rate their familiarity with the title characters (i.e.,

Peter Parker/Spider-Man, Thor, Steve Rogers/Captain America, and Bruce Banner/the Hulk) on a 5-point scale, 1 indicating little to no familiarity and 5 indicating extremely familiar. If participants had previously read the narratives included in this study, their responses were not included in analyses. Of the 63 participants recruited for this study, 4 indicated that they had read the stories before the study, and their responses were excluded from the analyses.

Table 3.6.

Participants' Indicated Familiarity with Characters

Characters	*M(SD)*
Peter Parker/Spider-Man	4.41(.83)
Steve Rogers/Captain America	3.51(1.25)
Bruce Banner/the Hulk	3.71(1.29)
Thor	3.53(1.45)

*Note: *N* = 51. Participants rated their familiarity on a 1-5 Likert scale. 1 = Not familiar at all. 2 = Slightly familiar. 3 = Somewhat familiar. 4 = Moderately familiar. 5 = Extremely familiar.

Finally, at the end of the session, participants were given a link in the Zoom chat feature and asked to click on the link and complete the demographic Qualtrics survey.

Scoring

Think-Aloud and Recall Coding Schemes

The goal of this study was to compare readers' online comprehension processes and offline comprehension products across text and non-linguistic graphic narratives. Therefore, I devised two coding schemes, one for participants' think-aloud responses to measure online processes and one for participants' recall responses to measure offline products (see Tables 7 and 8, respectively). Both coding schemes were based on theoretical frameworks for graphic narrative comprehension and previous think-aloud research (e.g., Cohn, 2020; Kendeou et al., 2020; Magliano et al., 2016; van den Broek et al., 2015, 2001).

Based on practice coding of two participants with incomplete data, five new think-aloud codes were added to the codes originally proposed for this study: internal state inferences; global and local inferences, character, scene, and affective responses (see Table 3.7). Five new codes were also added to the recall coding scheme originally proposed for this study: internal state inference (subcodes: valid/invalid); general inference (subcodes: valid/invalid), metacognitive comment, evaluation, and affective responses (see Table 3.8). These codes are also discussed in more detail in the *Think-aloud Coding* and *Recall Coding* sections below.

Think-aloud Coding. In order to compare the extent to which readers generate online processes across text and non-linguistic graphic narratives, the same think-aloud coding scheme was applied to responses for both types of modalities. First, to code participant verbal protocols from the think aloud tasks, responses to each sentence of text, as well as panel, was divided into idea units (i.e., typically subject-verb phrases), and then each idea unit was coded based on the coding scheme. The think-aloud coding scheme for both the text and graphic versions used in this study incorporated the iLC Framework (Kendeou et al., 2020), the PINS model (Cohn, 2020), and previous research with non-linguistic graphic and text narratives (e.g., Loughlin et al., 2015; Magliano & Graesser, 1991; Magliano et al., 2016; van den Broek et al., 2001; 2015).

Thus, responses from the think-aloud task were coded for *access* (i.e., descriptions of the visual or textual information; Cohn, 2020). For a response to be coded as *access*, the participant needed to not add any new information and simply describe or paraphrase what they saw or read. That is, in order for response to be coded as access, the participant must only be engaging in front-end processing (e.g., phonological, semantic, orthographic, lexical processing for text or scene gist, character/object, or motion processing for images) without yet engaging in back-end processing (e.g., inference generation; Magliano et al., 2013). In addition, the rest of the think-

aloud coding scheme include various types of inferences, narrative structure, and other cognitive processes.

For the *inferences*, I used the iLC Framework (Kendeou et al., 2020) and PINS Model (Cohn, 2020) as guides to code different types that readers generated during reading of text and non-linguistic graphic narratives (i.e., anaphoric, bridging, elaborative, emotion, predictions, updates). In addition, based on practice coding I conducted when developing my final coding scheme, I included *internal state* inferences to code responses about characters' thoughts and cognition that could not be described as emotion or otherwise. I also included *local* (i.e., references to the immediately preceding sentence or panel prior to the current sentence or panel) and *global* (i.e., references to previous sentences or panels not immediately preceding the current sentence or panel) inferences for responses that referred back to prior narrative information but did not fit the definition of anaphoric or bridging inferences. While anaphoric inferences refer back to a specific object or character previously introduced in the narrative and bridging inferences specifically infer an action not explicit depicted in the narrative, global or local inferences can more generally refer back to previous information (e.g., "She's watching TV again like she was in the beginning of the story.").

In addition, the PINS Model emphases processing of *narrative structure* while reading, so I included elements of narrative structure (i.e., goals, events, characters, scene, theme, author intent; Magliano & Graesser, 1991) as codes for all think-aloud responses. Codes included under *other processes* in the coding scheme were included as suggested by previous think-aloud studies with text and graphic narratives (e.g., Loughlin et al., 2015; Magliano & Graesser, 1991; Magliano et al., 2013; van den Broek et al., 2015; 2001). The think-aloud coding scheme and examples of responses are listed below in Table 3.7.

Pairs of four raters coded 10 of the participants' think-aloud responses in common (approximately 20%) out of 51 total. Specifically, the pairs of raters coded the think-aloud responses separately and then meet to discuss in common and discrepancies in coding. Interrater reliability for all of the think-aloud coding across the pairs of four raters was within the acceptable to good range (kappa = .71; Cohen, 1960).

Table 3.7.

Think-Aloud Coding Scheme

Category	Code	Definition	Examples
	Access	Description of visual/textual information; in graphic narrative, reader describes what they see; in text narrative, reader paraphrases or repeats text	*"I see a man and the scene is dark."* Text: She faced the alarm clock which read 5:29 A.M. Response: "*The alarm clock read 5:29."*
Inferences	Anaphoric Inferences	Determining a character or object is the same as referenced previously	*"Those must be the kids he saw in the photo earlier."*
	Bridging Inferences	Connect the current sentence/panel to the previous sentence/panel; Identify the action/event which connects the previous sentence/panel to the current sentence panel	Text: Mary Jane got up. Text: The tv played a commercial. Response: "*She turned on the TV."*
	Elaborative Inferences	Connecting current sentence/panel to background knowledge	*"I see a Christmas tree, so it must be the holiday season."*
	Emotion Inferences	Statement about the character's emotional state	*"She's sad." "He's feeling unsure."*
	Internal State Inferences	Statement about the character's internal state (e.g., thoughts) that are not emotions or goal-driven	*"She's thinking about Spider-Man."* *"She's wondering why he sat there."*
	Global/Local Inferences	Using information from previous panels to explain the current panel. Use Local for previous panel and Global for panels further away. Use only if anaphoric/bridging do not apply.	Text: Mary Jane modeled on the beach. Text: She moved down the red carpet. Response: "*Since she was modeling, she must be famous enough to go to a movie premier."*
	Predictions	Statement about what will occur next or later in the narrative	*"I think they're going to get caught."*
	Updates	Statement correcting or revising a previous statement	*"Oh, it's not a bus, it's a diner."*
Narrative Structure			

	Goals	Statement about the character's goals/objectives; could be super-ordinate or subordinate	*"He doesn't want his family to know." "He's trying to escape."*
	Events	Identification of narrative categories (e.g., introduction of conflict, climax, resolution, etc.); identification of beginning/end of events	*"Seems like we're setting up for the big conflict of the story."*
	Characters	Identification of who the main characters are or types of characters (e.g., hero, villain, protagonist, etc.)	*"This must be the main character."*
	Scene	Identification of the setting or change in scene	*"Now we're in a new place, looks like an office party."*
	Theme	Statements about the overall message or theme of the narrative	*"It's a story about forgiveness."*
	Author's intent	Statements about what the creators of the narrative are attempting to communicate to their audience	*"The writers are showing a contrast between when she's alone and when she's with people."*
Other	Metacognitive statements	Statements about agreement/disagreement and understanding/lack of understanding or background	*"I don't understand what's happening here."*
	Evaluations	Opinion statements about the content or structure of the narrative	*"Interesting."*
	Affective Responses	Statement about the participant's own personal emotional reaction to the story/panel/sentence.	*"Oh, that makes me sad for him."*
	Questions	Participant asks a question	*"Is that a bus?" "Why is he doing that?"*
	999	No response	

*Note: Internal State Inferences, Global/Local Inferences, Character, Setting, and Affective Responses were added to think-aloud coding scheme after practice coding. All codes applied to responses to both text and non-linguistic graphic narratives.

Recall Coding. In order to compare reader offline products across text and non-linguistic graphic narratives, the recall coding scheme also applied to responses for both types of modalities. Participants' recall responses were divided into idea units (i.e., typically subject-verb phrases), and then each idea unit was coded for recall response type. The recall coding scheme incorporated the PINS Model (Cohn, 2020) and previous research with non-linguistic graphic and text narratives (e.g., Magliano et al., 2016; van den Broek, 2005; van Neste et al., 2015). Previous studies involving recall tasks divide recall responses between explicit information stated in the story and reader generated or inferred information (e.g., Kendeou & van den Broek, 2005; van Neste et al., 2015). Therefore, the recall coding scheme for the proposed study was divided between participant responses stating an *explicit detail* from the narratives (i.e., story information, description) and statements not explicitly stated or shown in the narratives but that participants infer were coded as *inferences* (i.e., action, goals, emotions, internal states, general; see Table 3.8 for definitions and examples of each code). Based on coding schemes from previous works (Kendeou & van den Broek, 2005; van Neste et al., 2015), I also included an *Other* category which included metacognitive comments, evaluations, and affective responses.

Responses that recalled explicit content form the narratives were additionally coded as either *accurate* (i.e., reflective of information from the narrative) or *inaccurate* (i.e., included incorrect information from the narrative). For example, the response, "Mary Jane was alone in bed" would be coded as "Story Information – Accurate" because that information was included in the story. The response "Mary Jane went to the grocery store" would be coded as "Story Information – Inaccurate" because that event does not occur in the narrative. Likewise, responses that included inferences about the narrative were additionally coded as *valid* (i.e., supported by the narrative) or *invalid* (i.e., not supported by the narrative). For example, the response, "Bruce was

scared" would be coded as "Emotion – Valid" because the narrative was about Bruce hiding from FBI agents. The response, "Bruce wanted to fight the guys chasing him" would be coded as "Goal – Invalid" because based on the events of story, the character Bruce does not have the goal of wanting to fight his pursuers.

Pairs of four raters coded 10 of the participants' recall responses in common (approximately 21%) out of the total 48 participants. Pairs of raters coded the think-aloud responses separately and then meet to discuss in common and discrepancies in coding. Interrater reliability for the recall coding across the pairs of four raters was within the excellent range (kappa = .81; Cohen, 1960).

Table 3.8.

Recall Coding Scheme

Level of Representation	Category	Code	Definition	
Text / Image Base	Explicit	Story Information Subcode: Accurate/ inaccurate	Statement of something explicitly stated/depicted in the narrative	*"He walked into the restroom." "He placed his hand on his wife's shoulder."*
		Description Subcode: Accurate/ inaccurate	Description of a panel or sentence; focus is on text base/image rather than story content	*"It shows her crying." "You could see the background of a city." "It said he sipped his coffee."*
Situation Model	Inference	Action Subcode: valid/ invalid	An event not explicitly depicted in the story	*"He ordered coffee."*
		Goals Subcode: valid/ invalid	Statement of character's goals	*"He wanted his family to have a nice Christmas."*
		Emotion Subcode: valid/ invalid	Statement of character's emotional state	*"Even though she was famous, she was lonely."*
		Internal State Subcode: valid/ invalid	Statement of character's thoughts or cognition	*"She knew he was the Hulk." "He assumed he went through the window."*
		General Subcode: valid/ invalid	Statement of implicit content that is not an event/action, goal, or emotion	*"It was Christmas eve night."*
	Other	Metacognitive Comment	Statement about not remembering or not knowing	*"I can't remember what happened next." "The story didn't tell us what happened."*
		Evaluations	Opinion or judgement about the function or content of the story	*"It was nice that he still donated."*
		Affective Responses	Statement about the participant's own	*"I was glad the father wasn't upset with the boy."*

			emotional response to the story

*Note: Internal State, General, Metacognitive Comment, Evaluations, and Affective Responses were added to the recall coding scheme after pilot coding. All codes applied to responses to both text and non-linguistic graphic narratives.

Data Analyses

Data Entry and Computation

After coding think-aloud and recall responses, I calculated participant scores for each outcome variable. For online processing scores, I counted each time a code occurred in a participant's think-aloud responses for a narrative. I then added together the counts for the narratives a participant saw as text and for the narratives saw as non-linguistic graphic. Therefore, participants had a total count score for each code for text narratives and for non-linguistic graphic narratives. This score provided each participant a final score for each online process for both non-linguistic graphic and text. For example, every participant received a score for bridging inferences for text and a score for bridging inferences for non-linguistic graphic narratives. I made the same calculations for participants' recall responses to provide scores for each offline product measure. For the covariate measures, participants received a single total score for each measure, the ART and the VLF. I used these scores for my analyses.

Descriptive Statistics

I analyzed descriptive statistics for my data. These descriptive statistics included measures of normal distribution, including skewedness and kurtosis, as well as means and standard deviations. I conducted correlational analyzes among the variables in order to test assumptions of multivariate analyses (see *Multivariate Analysis of Covariance* below) and describe the relationship among the dependent variables.

Multivariate Analysis of Covariance

Because this study design involved two within groups (i.e., text and non-linguistic graphic) with multiple dependent measures and two covariates, I performed within groups, repeated measures multivariate analysis of covariance (MANCOVA; Huck, 2012) to answer my research questions. Two within-subjects, repeated measures MANCOVAs were conducted, one for the online comprehension process outcomes across text and non-linguistic graphic modalities and one for the offline comprehension product outcomes across text and non-linguistic graphic modalities. A repeated measures MANCOVA can indicate whether there was a significant difference across text and non-linguistic graphic processing or products. If the MANCOVA tests are significant, post hoc analyses (i.e., dependent t-tests to test mean differences across the two groups) can then be conducted to determine the source of the difference across the modalities. Thus, post hoc analyses in this study allow me to identify on which dependent variables readers differ across graphic and text reading.

MANCOVA Assumptions. A MANCOVA test requires a number of assumptions which I addressed before performing my analyses (Huck, 2012). The first assumption is independent random sampling. My sample was self-selected in that participants signed up for the study themselves through the SONA study system. Participants were then assigned a counterbalanced combination of the narratives they saw randomly. Thus, these two factors help address this first assumption. Second, MANCOVA requires independent variables to be categorical, while dependent variables are continuous. The independent variable for the current study, modality, is categorical, with the groups being either text or non-linguistic graphic. The dependent variables are continuous (e.g., number of bridging inferences generated, number of access statements generated). The covariates, participants' ART scores (Acheson et al., 2008; Stanovich & West, 1989) and VLF index (Cohn, 2014) scores, are also continuous. Thus, this second assumption has been met.

Third, MANCOVA assumes lack of multicollinearity among dependent variables. To test multicollinearity, I first created correlation matrices for the online process and offline product dependent variables. None of the variables correlated greater than .90; therefore, the assumption of lack of multicollinearity was met. Fourth, MANCOVA assumes multivariate normality. I assessed multivariate normality by examining skewedness and kurtosis. Because many of the dependent variables did violate standards of multivariate normality (see Tables 4.2 and 4.6 in the Results chapter), I used Greenhouse-Geisser corrections for the MANCOVA analyses which are robust to violations of normality (Abdi, 2010). Fifth, MANCOVA assumes homogeneity of variance. Because my analyses do not include between-subjects groups, tests of homogeneity are unnecessary. Finally, a MANCOVA assumes covariates are related to the dependent variables, and covariates should not strongly correlate with eacgh other. Through these correlational analyses, I found that ART and VLF index scores did correlate with each other. ART scores did not correlate with many of the dependent variables, and therefore, only VLF index scores were included as a covariate.

CHAPTER IV
RESULTS

Think-Aloud Responses

Descriptive Statistics. Table 4.1 shows the descriptive statistics for the screening and co-variate measures. Possible scores on the VLFI range from 0 to 52.5. Scores below 8 indicate low visual language fluency, scores from 10 to 19 indicate average fluency, and scores above 20 indicate high fluency (Cohn, 2014). This sample included participants who ranged from low to high frequency, with the mean near low fluency. Possible scores on the ART (Acheson et al., 2008) range from 0 to 65.

Table 4.1.

Descriptive Statistics for Screening and Covariate Measures for Think-Aloud Participant Sample

Measure	*M(SD)*	Min	Max
WJ-III Letter Word Identification	69.49(5.03)	59	78
ART	9.69(8.35)	0	37
VLFI	8.52(6.72)	1.50	32.13

Note: *N* = 51.

Descriptive statistics and indicators of normal distribution for think-aloud response variables are listed in Table 4.2. Acceptable levels of indicators of normal distribution or normality were set at skewedness scores of +/- 2 and kurtosis scores of +/- 7 (Kline, 2011). A number of online process variables from participants' think-aloud responses were outside acceptable levels of normality (see Table 4.2). For this reason, when conducting the MANCOVA for think-aloud responses, I will refer to the Greenhouse-Geisser correction because it is robust to violations of normality (Abdi, 2010).

Table 4.2.

Think-Aloud Descriptive and Distribution Statistics

		Non-Linguistic Graphic		Text	
		M(SD)	Skewedness/ Kurtosis	*M(SD)*	Skewedness/ Kurtosis
Inferences	Access	37.96(19.11)	1.06/1.43	18.31(16.28)	1.07/.15
	Anaphoric	2.04(2.18)	1.16/.98	.92(1.32)	2.41/6.14
	Bridging	5.75(3.73)	.51/-.13	4.45(3.67)	1.28/.87
	Elaboration	45.25(16.96)	.39/-.16	44.84(18.87)	.37/-.00
	Emotion	13.47(8.46)	.94/.40	10.12(5.47)	1.08/1.68
	Internal State	7.45(3.93)	.49/-.34	7.63(5.15)	1.87/4.96
	Global	3.76(3.02)	1.29/2.71	4.90(4.32)	1.84/5.31
	Local	2.63(3.58)	3.94/20.93	3.25(3.48)	1.61/2.81
	Prediction	4.49(3.87)	1.77/4.96	6.80(4.68)	.49/-.66
	Update	2.86(2.98)	1.18/.88	2.78(3.00)	1.19/.29
Narrative	Goal	8.24(5.50)	1.04/.35	11.75(6.33)	.75/.34
	Event	.08(.27)	3.23/8.79	.04(.20)	4.89/22.83
	Character	.31(.86)	3.06/9.10	.31(.68)	2.33/5.17
	Scene	.65(1.15)	3.23/12.41	.51(.86)	1.75/2.37
	Theme	.08(.34)	4.69/23.10	.27(.72)	2.50/5.05
Others	Author Intent	.24(.74)	3.96/16.75	.27(.78)	3.47/12.82
	Metacognition	3.71(5.50)	2.90/10.40	3.20(4.85)	2.77/8.89
	Evaluation	2.33(3.27)	2.95/12.74	2.25(3.70)	2.02/3.47
	Affective Response	.22(.64)	3.08/8.92	.76(2.06)	3.86/15.40
	Question	.80(2.47)	5.88/38.32	.86(1.93)	2.60/6.17

Note: N = 51. Scores are based on the total counts participants generated across the two non-lingusitic graphic narratives and the two text narratives.

Correlations. Correlational analyses were conducted to determine lack of multicollinearity among the dependent variables and that covariate measures correlated with dependent variables. That is, a strong correlation between dependent variables would indicate that the two variables are measuring the same construct. Correlational relationships between online process variables from think-aloud responses are listed in an external spreadsheet (https://bit.ly/ness-maddox_dissertation_correlations). The focus of this study first research question is to investigate differences in processes across text and non-linguistic graphic narratives, not to investigate the relationships across processes within text narrative reading or within non-linguistic graphic narrative reading. Therefore, I highlight only significant relationships across text and non-linguistic graphic narratives. Regarding inferences, the correlation analyses showed significant relationships for text and non-linguistic graphic anaphoric inferences ($r = .50$, $p < .001$), text and non-linguistic graphic bridging inferences ($r = .36$, $p < .01$), text and non-linguistic graphic elaborations ($r = .83$, $p < .001$), and text and non-linguistic graphic internal state inferences ($r = .034$, $p < .05$). Of all the types of inferences generated during think-aloud tasks, emotion inferences were not significantly related across text and non-linguistic graphic narratives. Access statements were also significantly correlated across text and non-linguistic graphic narratives ($r = .44$, $p < .001$), as well as updates ($r = .059$, $p < .001$). Predictions were not significantly related across text and non-linguistic graphic narratives. Of statements about narrative structure, only statements about events ($r = .32$, $p < .05$) and characters ($r = .38$, $p < .01$) were related across text and non-linguistic graphic narratives. Finally, statements about readers' metacognition ($r = .84$, $p < .001$), evaluations ($r = .56$, $p < .001$), affective responses ($r = .65$, $p < .001$), and questions ($r = .59$, $p < .001$) were related across text and non-linguistic graphic narratives.

Covariates for MANCOVAs should correlate with the dependent variables but not strongly correlate with each other (Huck, 2012). ART scores correlated with VLF index scores ($r = .51$, $p < .01$) text elaborations ($r = .34$, $p < .05$), local inferences ($r = .32$, $p < .05$), and theme, ($r = .31$, $p < .05$), as well as non-linguistic graphic events ($r = .31$, $p < .05$), and affective response ($r = .40$, $p < .01$). VLF index scores correlated with text elaborations ($r = .35$, $p < .05$), scene ($r = .29$, $p < .05$), theme ($r = .46$, $p < .01$), author intention ($r = .49$, $p < .01$), and affective response ($r = .40$, $p < .01$), as well as non-linguistic graphic elaborations ($r = .28$, $p < .05$), scene ($r = .42$, $p < .01$), author intention ($r = .52$, $p < .01$), metacognition ($r = .30$, $p < .05$), and affective response ($r = .43$, $p < .01$). Since ART scores significantly correlated with VLF index scores, and VLF index scores correlated with more think-aloud response variables, I included only included VLFI scores as a covariate measure in the MANCOVA for recall responses. The strongest relationship among think-aloud responses variables was statements about author intentions across text and non-linguistic graphic narratives ($r = .86$); therefore, the assumption of non-multicollinearity is not violated.

MANCOVA. Because the sample included within-subjects responses and variables that violated assumption of normality, I referred to Greenhouse-Geisser correction statistics for tests of within-subjects effects (see Table 4.3). Effects of modality were significant ($F(1, 51) = 13.523$, $p < .001$, $\eta^2 = .216$) with a large effect size. Overall, participants generated more response statements for non-linguistic graphic narratives than text. Effects of process were also significant ($F(19, 51) = 71.521$, $p < .001$, $\eta^2 = .593$) with a large effect size. For example, participants generated more elaboration response statements than other type of process response statements. The interaction of modality and process is significant ($F(19, 51) = 9.188$, $p < .001$, $\eta^2 = .159$) with a large effect size. This finding indicates that participants did differ in the extent to

which they generated online comprehension processes across modalities, supporting the main hypothesis. The interaction between process and VLF index scores is also significant ($F(19, 51) = 3.623$, $p = .017$, $\eta^2 = .069$) with a medium effect size. The interaction between modality, process, and VLF index scores is non-significant ($F(19, 51) = .673$, $p = .587$).

Because the interaction between modality and process was significant, I conducted post hoc analyses to determine the source of the difference between modalities.

Table 4.3.

Tests of Within-Subjects Effects for Think-Aloud Responses

	DF	*F*	*p*-value	η^2	1 – β
Modality	1	13.523	<.001	.216	.950
Modality*VLFI	1	.589	.447	.012	.117
Process	19	71.521	<.001	.593	1.00
Modality*Process	19	9.188	<.001	.158	.998
Process*VLFI	19	3.623	.017	.069	.795
Modality*Process*VLFI	19	.673	.587	.014	.200

Note: All *F* statistics are from Greenhouse-Geisser tests. 1 – β is the observed power. *N* = 51.

Post Hoc Analyses. Post hoc analyses involved dependent t-tests comparing process variables coded from think-aloud responses between non-linguistic graphic and text narratives (see Table 4.3 and Figure 4.1). I hypothesized that participants would generate more anaphoric, bridging, elaborative, emotion, global/local, predictive, and updating inferences for non-linguistic graphic narratives than text narratives (e.g., Coderre, 2020; Cohn & Magliano, 2020; Gnepp, 1983; Loughlin et al., 2015). I also hypothesized that participants would generate more access statements and more statements about narrative events and scene for non-linguistic graphic narratives than text narratives. On the other hand, I expected that participants would generate more internal state inferences and goal statements for text than non-linguistic graphic narratives (e.g., Cohn & Magliano, 2020; Gnepp, 1983; Loughlin et al., 2015). Results of post hoc analyses support that differences between modalities depends on the process. As expected, participants generated more access statements for non-linguistic graphic narratives (*M* = 37.96) than text narratives (*M* = 18.31). They also generated more bridging inferences for non-linguistic graphic (*M* = 5.75) than text narratives (*M* = 4.45), more anaphoric inferences for non-linguistic graphic (*M* = 2.04) than text narratives (*M* = .92), and more emotion inferences for non-linguistic graphic (*M* = 13.47) than text narratives (*M* = 10.12).

For some processes, participants generated more responses for text narratives than non-linguistic graphic narratives. As expected, participants generated more goal statements for text narratives ($M = 11.75$) than non-linguistic graphic narratives ($M = 8.24$). Additionally, participants generated more predictions for text narratives ($M = 6.80$) than non-linguistic graphic narratives ($M = 4.49$), which was not expected. Contrary to my expectations, participants did not differ between text and non-linguistic graphic narratives for elaborations, internal state inferences, global and local, other narrative structure categories. Additionally, they generated more affective responses for text narratives ($M = .76$) than non-linguistic graphic narratives ($M = .22$).

Table 4.4.

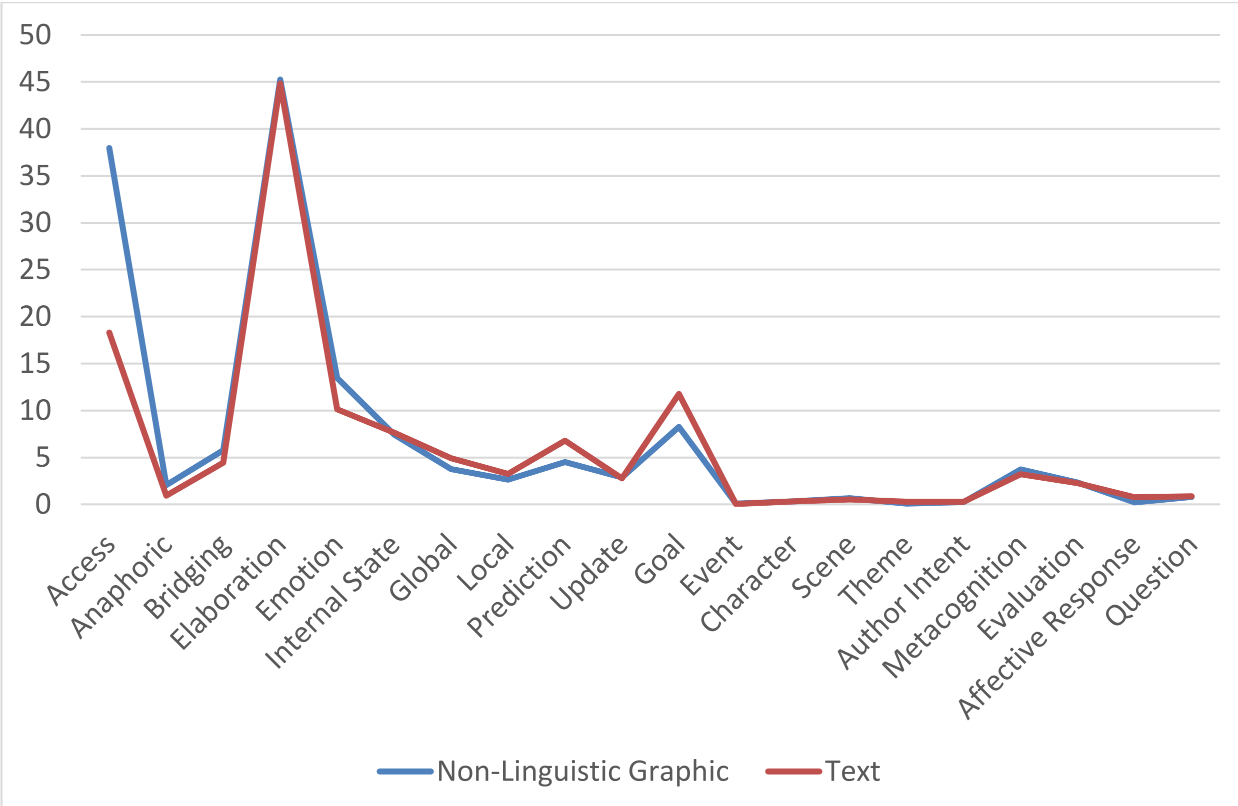
Post Hoc Analyses for Process Variables from Think-Aloud Responses

Process	t-value	2-tailed p-value	Cohen's d Effect size	95% Confidence Interval Lower / Upper
Access	-7.434	<.001***	-1.041	-1.389 / -.696
Anaphoric	-4.192	<.001***	-.587	-.882 / -.287
Bridging	-2.200	.032*	-.308	-.588 / -.026
Elaboration	-.276	.784	-.039	-.313 / .236
Emotion	-2.440	.018*	-.342	-.622 / -.058
Internal State	.237	.814	.033	-.242 / .308
Global	1.654	.104	.232	-.048 / .509
Local	1.007	.319	.141	-.136 / .416
Prediction	3.004	.004**	.140	.132 / .705
Updates	-.207	.837	-.029	-.303 / .244
Goals	3.309	.002**	.463	.172 / .750
Events	-1.000	.322	-.140	-.415 / .136
Character	.000	1.00	.000	-.274 / .274
Scene	-.765	.448	.238	-.382 / .169
Theme	1.697	.096	.238	-.042 / .515
Author Intent	.704	.485	.099	-.177 / .373
Metacognition	-1.223	.227	-.171	-.447 / .016
Evaluation	-.170	.865	-.024	-.296 / .251
Affective Response	2.29	.026*	.321	.038 / .601
Questions	.204	.839	.029	-.246 / .303

Note: * indicates significant at the .05 level, ** is significant at the .01, and *** is significant at the .001. $N = 51$.

Figure 4.1.

Means of Online Processes Across Non-Linguistic Graphic and Text Narratives



Recall Responses

Missing Data. Of the 51 participants whose think-aloud responses were analyzed, three of those participants did not provide recall responses to at least one of the narratives. Thus, the sample size for recall data is 48 participants.

Descriptive Statistics. Table 4.5 shows the descriptive statistics for the screening and co-variate measures. Possible scores on the VLFI range from 0 to 52.5. Scores below 8 indicate low visual language fluency, scores from 10 to 19 indicate average fluency, and scores above 20 indicate high fluency (Cohn, 2014). This sample included participants who ranged from low to high frequency, with the mean near low fluency. Possible scores on the ART (Acheson et al., 2008) range from 0 to 65.

Table 4.5.

Descriptive Statistics for Screening and Covariate Measures for Recall Participant Sample

Measure	*M(SD)*	Min	Max
WJ-III Letter Word Identification	69.56(5.14)	59	78
ART	9.69(8.51)	0	37
VLFI	8.58(6.81)	1.50	32.13

Note: *N* = 48.

Descriptive statistics and indicators of normal distribution for recall response variables are listed in Table 4.6. Acceptable levels of indicators of normal distribution or normality were set at skewedness scores of +/- 2 and kurtosis scores of +/- 7 (Kline, 2011). Many of the product variables from recall responses were outside the acceptable levels of normality. For this reason, when conducting the MANCOVA for recall responses, I will refer to the Greenhouse-Geisser correction because it is robust to violations of normality (Abdi, 2010).

Table 4.6.

Descriptive and Distribution Statistics for Product Variables from Recall Responses

			Non-Linguistic Graphic		Text	
			M(SD)	Skewedness/ Kurtosis	*M(SD)*	Skewedness/ Kurtosis
Text / Image Base	Explicit	Description				
		Accurate	1.50(2.79)	3.26/12.47	1.42(1.93)	1.63/2.54
		Inaccurate	.10(.47)	5.41/31.63	.02(.14)	6.93/48.00
		Story Information				
		Accurate	17.69(9.10)	1.11/1.78	25.77(12.47)	.37/-.20
		Inaccurate	1.12(1.47)	2.39/9.02	1.42(1.72)	1.11/.13
Situation Model	Inferences	Action				
		Valid	12.46(5.27)	.36/-.38	12.65(5.96)	.15/-.53
		Invalid	.65(1.04)	1.60/1.75	.33(.81)	3.10/10.48
		Goals				
		Valid	3.52(2.69)	2.24/8.69	3.10(2.72)	1.28/1.51
		Invalid	.06(.25)	3.73/12.45	.02(.14)	6.93/48.00
		Emotion				
		Valid	4.04(2.59)	.53/-.63	2.85(2.42)	.96/.57
		Invalid	.04(.20)	4.74/21.32	.06(.25)	3.73/12.45
		Internal State				
		Valid	3.13(2.49)	.92/.22	2.81(2.46)	.99/.45
		Invalid	.15(.36)	2.07/2.39	--*	--/--*
		General				
		Valid	5.13(3.76)	1.52/2.89	4.21(3.38)	.698 / -.521
		Invalid	.13(.393)	3.37/11.75	.04(.202)	4.737/21.323
	Others	Metacognition	.50(.88)	2.19/5.33	.96(1.77)	2.36/5.60
		Evaluation	.35(.53)	1.08/.11	.23(.52)	2.25/4.43
		Affective Response	--*	--*	.12(.53)	4.60/21.52

Note: *Participants did not include any invalid internal state inferences for text narratives or any affective responses for graphic narratives. *N* = 48. Scores are based on the total counts participants generated across the two non-linguistic graphic narratives and the two text narratives.

Correlations. Correlational analyses were conducted to determine lack of multicollinearity among the dependent variables and that covariate measures correlated with dependent variables. Correlational relationships for product variables from recall responses are shared in an external spreadsheet (see https://bit.ly/ness-maddox_dissertation_correlations). The focus of this study's second research question is to investigate differences in offline products across text and non-linguistic graphic narratives and not relationships across comprehension products within recall responses for text narratives or within recall responses for graphic narratives. Therefore, I highlight only significant relationships across text and non-linguistic graphic narratives. Regarding participants' recall of explicit information in the narratives, accurate information ($r = .68$, $p < .001$) and accurate description ($r = .32$, $p < .05$) significantly correlated across text and non-linguistic graphic narratives. For inferences, valid action inferences ($r = .56$, $p < .001$), invalid emotion inferences ($r = .34$, $p < .01$), and valid general inferences ($r = .45$, $p < .001$) significantly correlated across text and non-linguistic graphic narratives. Finally, statements about metacognition correlated across text and non-linguistic graphic narratives ($r = .49$, $p < .001$).

Covariates for MANCOVAs should correlate with the dependent variables (Huck, 2012). ART scores significantly correlated with VLF index scores ($r = .51$, p < .01) and evaluations for non-linguistic graphic narratives ($r = .43$, p < .01). VLF index scores significantly correlated with text valid goal inferences ($r = .36$, $p < .05$), valid internal state inferences ($r = .32$, $p < .05$), and metacognition ($r = .35$, $p < .05$), as well as non-linguistic valid action ($r = .33$, $p < .05$), valid general inferences ($r = .38$, $p < .01$), and evaluations ($r = .29$, $p < .05$). Since ART scores did not correlate with most recall response variables but was significantly correlated with VLFI scores which correlated with more of the recall response variables, I included only included VLFI scores as a covariate measure in the MANCOVA for recall responses. The strongest relationship

among recall responses variables accurate information across text and non-linguistic graphic narratives (*r* = .68); therefore, the assumption of non-multicollinearity is not violated

MANCOVA. Because the sample included within-subjects responses, I referred to Greenhouse-Geisser correction statistics for tests of within-subjects effects (see Table 4.7). Because participants did not produce invalid internal state inferences for text narratives or affective responses for text narratives, those variables were not included in tests of within-subjects effects. Effects of modality were non-significant ($F(1, 48) = .164$, $p = .687$). That is, participants did not recall more product variables overall for one modality. Effects of the interaction between modality and VLF index scores was significant ($F(1, 48) = 6.050$, $p < .018$, $\eta^2 = .116$), with a medium effect size. Effects of product were significant ($F(15, 48) = 54.752$, $p < .001$, $\eta^2 = .543$) with a large effect size. For example, participants recalled more accurate story information than other product variables. Effects of the interaction between modality and product variables ($F(15, 48) = 3.698$, $p = .010$, $\eta^2 = .074$) with a medium effect size. This finding indicates that participants did differ in comprehension products across modality, supporting the main hypothesis. Effects of the interaction between product variables and VLF index scores were non-significant ($F(15, 48) = 2.535$, p = .092). Effects of the interaction between modality, products, and VLF index scores were also not significant ($F(15, 48) = 1.943$, $p = .116$). Because effects of the interaction between modality and product variables was significant, post hoc analyses are warranted to investigate the source of the differences between modalities.

Table 4.7.

Tests of Within-Subjects Effects for Recall Responses

	DF	*F*	*p*-value	η^2	1 – β
Modality	1	.164	.687	.004	.068
Modality*VLFI	1	6.050	.018	.116	.673
Product	15	54.752	<.001	.543	1.00
Modality*Product	15	3.698	.10	.074	.836
Product*VLFI	15	2.535	.092	.052	.463
Modality*Product*VLFI	15	1.943	.116	.041	.532

Note: All *F* statistics are from Greenhouse-Geisser tests. 1 – β is the observed power. *N* = 48.

Post Hoc Analyses. Post hoc analyses involved dependent t-tests to compare product variables from recall responses between non-linguistic graphic and text narratives (see Table 4.7 and Figure 4.2).

I expected that regarding participants' text/image base level of representation for the narratives, participants would include more accurate story information and more accurate description statements in recall of non-linguistic graphic narratives than text narratives (Loughlin et al., 2015; Magliano et al., 2013). Contrary to these expectations, participants included more accurate information in their recall responses for text narratives (*M* = 25.77) than non-linguistic graphic narratives (*M* = 17.69). Participants did not differ in accurate description statements between modalities.

Regarding participants' situation model level of representation, I expected that participants would include more valid inferences in recall of non-linguistic graphic narratives than text narratives (Cohn, 2013; Cutting, 2016; Magliano et al., 2013). Again, I expected two exceptions; that is, that participants would include more valid goal inferences and valid internal state inferences in recall of text narratives than in recall of non-linguistic graphic narratives (Johnson & Mandler; 1980; Magliano et al., 2013; Rumelhart, 1977; Stein & Glenn,1979). Contrary to these

hypotheses, overall, participants did not differ in the inclusion of inferences during recall across text and non-linguistic graphic narratives. However, participants did generate more valid emotion inferences in their recall responses for non-linguistic graphic narratives ($M = 4.04$) than text narratives ($M = 2.85$), as expected. Additionally, participants included more metacognitive statements in their recall responses for text narratives ($M = .96$) than non-linguistic graphic narratives ($M = .50$). These results and how they relate to my hypotheses are discussed more thoroughly in the following chapter, Chapter V: Discussion.

Table 4.8.

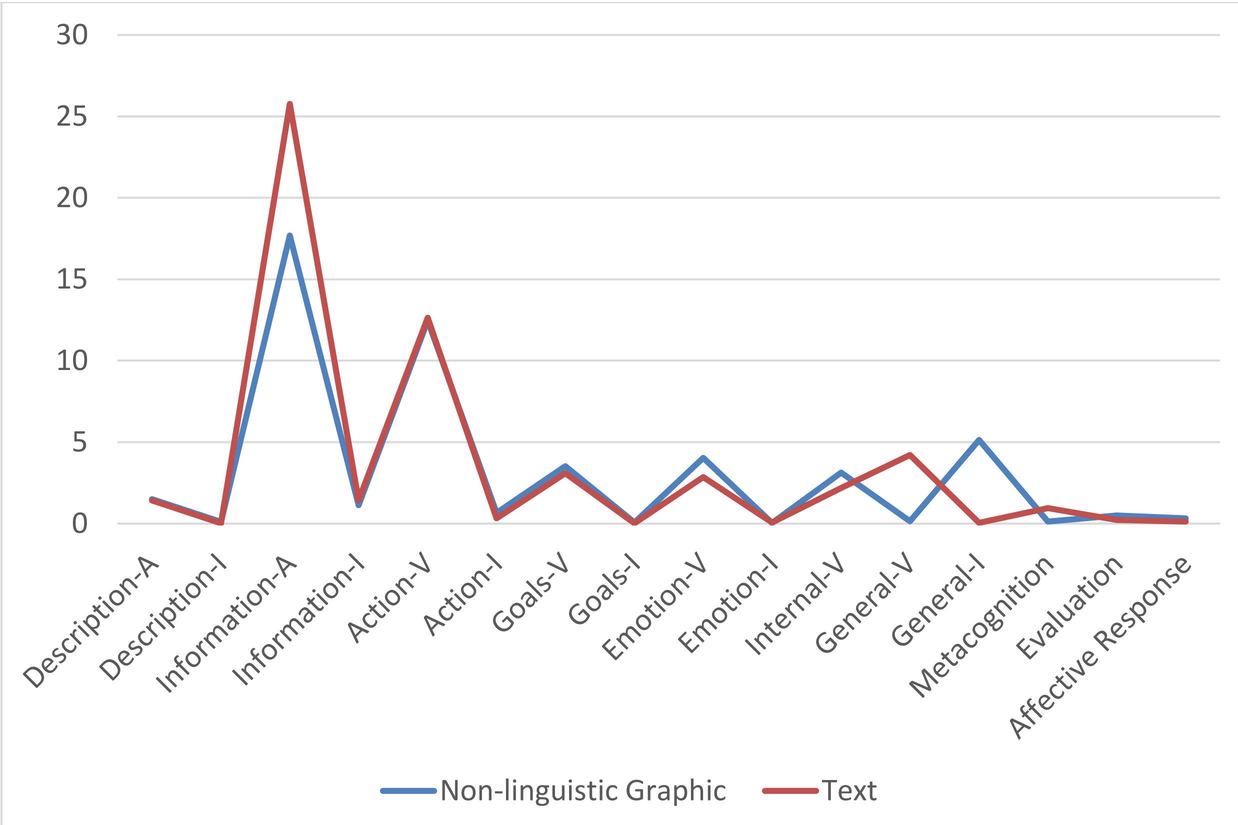
Post Hoc Analyses of Product Variables from Recall Responses

Level of Representation		Product Variable		t-value	p-value	Cohen's d Effect Size	Confidence Interval Lower/Upper
Text / Image Base	Explicit	Description					
			Accurate	-.203	.840	-.029	-.312/.254
			Inaccurate	-1.159	.255	-.167	-.451/.119
		Story Information					
			Accurate	6.083	<.001***	.878	.541/1.208
			Inaccurate	1.050	.299	.152	-.134/.435
Situation Model	Inferences	Action					
			Valid	.245	.807	.035	-.248/.318
			Invalid	-1.796	.079	-.259	-.546/.030
		Goals					
			Valid	-.841	.405	-.121	-.405/.163
			Invalid	-1.000	.322	-.144	-.428/.141
		Emotion					
			Valid	-2.300	.026*	-.333	-.621/-.040
			Invalid	.573	.569	.083	-.201/.366
		Internal State					
			Valid	-.677	.502	-.098	-.381/.186
		General					
			Valid	-1.685	.099	-.243	-.529/.045
			Invalid	-1.430	.159	-.206	-.491/.081
	Other	Metacognition		2.058	.045*	.297	.006/.585
		Evaluation		-1.288	.204	-.186	-.470/.100

Note: * indicates significant at the .05 level, ** is significant at the .01, and *** is significant at the .001. $N = 48$

Figure 4.2.

Means of Offline Products Across Non-Linguistic and Text Narratives



CHAPTER V
DISCUSSION

Readers generate comprehension processes while reading narratives (i.e., online) or comprehension products (e.g., text /image base, situation model) after reading (i.e., offline). Previous research has not investigated whether linguistic information in text narratives and visual information in non-linguistic graphic narratives influence readers' generation of online processes or offline products. The purpose of this study was to compare online processes generated while reading and offline products readers include during recall after reading across text and non-linguistic graphic narratives. Previous research has shown that a reader's experience and exposure to both text and graphic narratives influences their comprehension of a narrative of that modality. I investigated whether text and non-linguistic graphic modalities had an influence on readers' online processing and offline products beyond what experience and exposure explained. Therefore, my research questions were 1) to what extent do readers differ in the generation of comprehension processes while reading text and non-linguistic graphic narratives after controlling for text print exposure and visual language fluency, and 2) to what extent do readers differ in their inclusion of comprehension products for text and non-linguistic graphic narratives after controlling for text print exposure and visual language fluency?

In order to answer those research questions, I conducted a study in which participants read both text and non-linguistic graphic narratives. For every story, participants thought aloud while reading, and after reading, they recalled the narrative just read. Participants were also given measures of text print exposure, the Author Recognition Test (Acheson et al., 2008), and experience with graphic narratives, the Visual Language Fluency Index (Cohn, 2014). My participants consisted of adult readers who ranged from low to high on measures of print exposure and visual language fluency.

After recording participants' think-aloud and recall responses, I coded their responses based on processes and products identified by my theoretical foundation and previous research with text. Think-aloud responses were coded for access, bridging inferences, anaphoric inferences, emotion inferences, internal state inferences, global and local inferences, predictions, and updates. Think-aloud responses were also coded for goal statements and other statements about the narrative (i.e., events, scene, and characters). They were also coded for statements about theme and author intent and readers' metacognition, evaluations, affective response, and questions. Recall responses were coded for accurate and inaccurate story information and description. They were also coded for valid and invalid inferences, including action, goals, emotion, internal states, and general inferences. Finally, as with the think-aloud responses, recall responses were coded for readers' metacognition, evaluations, and affective responses. I expected that participants would differ in the extent to which they generate processes for text and non-linguistic graphic narratives during the think-aloud tasks as well as the extent to which they included product variables in their text/image base and situation model representations during recall.

Based on the MANCOVAs conducted, tests of within subjects effects were significant, indicating differences in processing between text and non-linguistic graphic narratives as well as differences in products between text and non-linguistic graphic narratives. Post hoc analyses indicate that differences between the two modalities depends on the process or product variables examined and that these processes or products of comprehension are not always greater for non-linguistic graphic narratives, as the Visual Ease Assumption would expect (see Coderre, 2020; Cohn & Magliano, 2020). I provide detail for differences in processing during the think-aloud tasks and products from the recall tasks in the following sections.

Differences in Processing from Think-Aloud Responses

I hypothesized that results would indicate that participants would generate online comprehension processes to different extents as observed in think-aloud tasks across text and non-linguistic graphic narratives after controlling for participants' text print exposure and visual language fluency. Because text print exposure did not correlate with many of the processes I investigated, I only controlled for visual language fluency in my analyses. Overall, I expected that participants would generate more anaphoric, bridging, elaborative, emotion, global/local, predictive, and updating inferences for non-linguistic graphic narratives than text narratives (e.g., Coderre, 2020; Cohn, 2013; Cohn & Magliano, 2020; Cutting, 2016; Gnepp, 1983; Loughlin et al., 2015; Magliano et al., 2013). Additionally, I expected that participants would include more access statements as well as narrative event and scene statements. However, I also expected that participants would generate more internal state inferences and goal statements during reading of text narratives than non-linguistic graphic narratives (e.g., Cohn & Magliano et al., 2020; Johnson & Mandler; 1980; Magliano et al., 2013; Magliano et al., 2018; Rumelhart, 1977; Stein & Glenn,1979).

In general, participants produced more idea units in their think-aloud responses for non-linguistic graphic narratives than for text narratives. That is, participants, overall, produced more responses for non-linguistic graphic narratives than text. Post hoc analyses revealed that participants generated more access statements as well as anaphoric, bridging, and emotion inferences for non-linguistic graphic narratives and generated more predictions and goal statements for text narratives. While these results did not support all of my expectations for participants' online processing, the results do support how participants processed text and non-linguistic graphic narratives differently depending on the process examined. Next, I discuss the processes in which participants differed between text and non-linguistic graphic narratives.

Access

Post hoc analyses indicated that participants generated more access responses for non-linguistic graphic narratives than text narratives. For non-linguistic graphic narratives, access responses involved describing pictures or stating what is happening in the panel without the inclusion of background knowledge or connecting to other panels. For text narratives, access responses were similar to paraphrasing; that is, participants repeated or rephrased the text in their own words without including background knowledge or connecting to other sentences. Thus, for either narrative modality, access responses involved participants taking in and processing the surface form and text/image base of the panel or sentence they were currently reading (Cohn, 2020; Kintsch, 1988). These post hoc results indicate that participants generated more idea units describing individual panels than they did repeating or rephasing text. Results indicated a large effect of modality on access responses. This difference between modalities is perhaps due to the adage, “A picture is worth a thousand words.” When paraphrasing text, a reader can only repeat or rephase words for so long, but a viewer of an image could perhaps find many elements to describe. In which case, readers of graphic narratives have more elements to integrate into their image base representation than when reading text and constructing a text base representation. This finding may illuminate how readers need to expend more cognitive effort to construct the image base for non-linguistic graphic narratives than the text base of text narratives because graphic panels present more elements than a single sentence (McCloud, 1993).

For example, consider the panel in Figure 5.1, the first panel from Narrative D. The corresponding sentence from the text version of this narrative says, “Young Thor and Loki played in the castle garden with wooden swords, Thor’s sword broken from Loki’s swing.” To paraphrase or produce access responses to the sentence version of this part of the story, a participant may

say, “Ok, Thor and Loki were playing in the gardens. They had wooden swords. Thor’s was broken.” A participant viewing the panel, however, could comment on the bird bath, the stone trail beside Thor and Loki, the colors of their outfits, the framing of the image, and so on. Thus, a reader has more information to access from a visual image as opposed to a single text sentence, whether or not that visual information is relevant to the story.

These findings extend the current literature because previously no study had directly compared the processes used to develop the text/image base representations across text and non-linguistic graphic narratives during online processing. Previous research has compared the text base representation across modalities, but these modalities included linguistic information and tested readers’ text base representations after reading rather than during reading. For example, Wannagat et al. (2017) compared child and adult readers’ text base representations for audiovisual, auditory, and written text presentations. They found that adult and 10-year-old readers do not differ in their text base representations across modalities, but visual information benefits 8-year-old readers’ text base representations. Wannagat and colleague’s study compared participants’ text base representations for linguistic information when presented along with visual information. The current study extends the literature by directly comparing participants’ processes generated to develop the text base representations for text narratives with the construction of their image base representations for non-linguistic graphic narratives by accessing the linguistic and visual information presented in the narratives.

Figure 5.1.

Panel from Narrative D

Note: Image is from *Thor* volume 2, issue 44 (Jurgens et al., 2002).

Anaphoric Inferences

Post hoc analyses also indicated that participants generated more anaphoric inferences for non-linguistic graphic narrative than text narratives. Results indicated a medium effect of modality on anaphoric inferences. Anaphoric inferences involve participants connecting current information to previous information in the narrative, specifically that an object, character, or place is the same as introduced previously in the narrative (Coopmans & Cohn, 2022; Kendeou et al., 2020). The connection to previous objects, characters, and places may be more obvious to text readers and do not require a verbal inference because those words are stated explicitly in text. For example, participants saw either the non-linguistic graphic version or the text version of a narrative about a man going home after work. For the non-linguistic graphic version, participants might comment, “That’s the same man from the last panel.” In the text version, the man is referred to as “Lloyd,” and participants did not state “That’s the same Lloyd from the last sentence.” Seeing a word for an object or character repeated in the text may be enough information for readers to know it is the same object or character previously introduced in the narrative, but

seeing the image of an object or character still may encourage a reader to generate that connection. This difference in anaphoric inference may indicate that readers may integrate information into their text base representation for text but need to generate inferences to integrate the same information into their situation model representation for graphic narratives.

These results support previous research that investigated anaphoric inferences generated during graphic narrative reading which found that non-linguistic graphic readers do generate anaphoric inferences to recognize current characters and objects as previously seen characters and objects (Coopmans & Cohn, 2022). Participants in the current study did generate anaphoric inferences for both text and non-linguistic graphic narratives. No previous study had directly compared the generation of anaphoric inferences across text and non-linguistic graphic narratives; thus, this study extends previous research in this area. However, given that this has been the first study to directly examine this inference type by text and graphic narratives, additional research should be conducted to replicate these findings.

Bridging Inferences

Post hoc results also indicated that participants generated more bridging inferences for non-linguistic graphic narrative than text narratives. These results were qualified by a small effect of modality on bridging inferences. Bridging inferences, like anaphoric inferences, also connect to previous information in the text or graphic narrative, in which participants fill in actions not explicitly depicted or stated in the narrative and connect ideas across panels or sentences (Cohn, 2020; Cohn & Wittenburg, 2015; Huff et al., 2020; Magliano et al., 2016; Manfredi et al., 2017). Generating bridging inferences is necessary to comprehend both text and graphic narratives but features of graphic narratives may prompt readers to generate more bridging inferences.

For instance, in a typical graphic narrative with multiple panels on each page, panels are separated by blank space, called the “gutter” (McCloud, 1993). Graphic narrative readers infer the actions and events that must happen in the blank space of the gutter between panels (Cohn & Wittenburg, 2015; Manfredi et al., 2017), especially when narrative context supports the generation of a bridging inference (Cohn, 2021; Cohn & Kutas, 2015). A reader could simply view each panel as a standalone image, but by generating bridging inferences, graphic readers connect these separate panels into a coherent narrative in their mental representation. Because text sentences overlap in explicit words minimizing the coherence gap across sentences (McNamara & Kintsch, 1996; van Dijk & Kintsch, 1983), text readers may not need to generate as many bridging inferences to create a coherent mental representation of the text as they do during graphic narrative comprehension (Trabasso et al., 1989; Trabasso & Magliano, 1996).

The current study extends previous findings by comparing the generation of bridging inferences for an all text narrative and an all visual non-linguistic narrative. Previous research has compared the generation of a bridging inference at specific points in the narrative when a bridging event is removed (e.g., Magliano et al.., 2016) or when the narrative contains both text and visual information (e.g., Huff et al., 2020). Huff et al. (2020) compared bridging inferences between text and graphic information, but participants were required to switch between modalities within a narrative. Huff et al.’s results support that switching between modalities increased difficulty for generating bridging inferences for text. Participants in the current study did switch between modalities in that they read either a text or non-linguistic graphic narrative and then switched to the other modality for the next narrative and so on. They did not switch modalities within a narrative. Thus, results of the current study do not provide any evidence for how switch-

ing between modalities for each narrative affected participants' ability to generate bridging inferences for text; rather, the differences between modalities can be attributed either text or graphic narrative. Participants did generate bridging inferences for text narratives, just to a lesser extent than for non-linguistic graphic narratives.

Emotion Inferences

Post hoc analyses also indicated that participants generated more emotion inferences for non-linguistic graphic narrative than text narratives. These results were qualified by a small effect of modality on emotion inferences. During the think-aloud task, emotion inferences involved statements about the characters' emotional states or moods. In the text versions of the narrative, emotional words were never included, only descriptions of facial expressions. In the non-linguistic graphic narrative versions, faces are depicted more than facial expressions are described in text. Thus, readers of graphic narratives may have been drawn to attend to the faces depicted in the images (Laubrock et al., 2018), and participants in this study may have generated emotion inferences based on such facial expressions seen. These findings support previous research which has identified that seeing faces prompts people to generate inferences about emotions (Ekman, 1992; Ness-Maddox et al., 2022), even if the facial expression is not the main focus of the panel. Additionally, previous research has shown that as children develop their socioemotional knowledge (i.e., 6-years-old), they integrate information from characters' facial expressions and situational cues; that is, they generate nuanced emotional inferences based on two sources of visual information, the character and the scene (e.g., Gnepp, 1983). In the current study, participants had the benefit of scene and setting information in non-linguistic graphic narratives to guide emotional inferences (Gnepp, 1983).

For example, consider the panel in Figure 5.2 from Narrative B. The corresponding sentence for this point in the story is, “Lloyd walked down the street while festive people surged by many carrying packages.” From the panel, a reader sees the smiles on people’s faces, while Lloyd hunched and looking down. The reader may infer from Lloyd’s body language that he is sad, while everyone around him is smiling and happy. The shading of this panel with Lloyd lighted and everyone around him in shadows also may imply to readers that Lloyd feels alone in a crowd of people. Thus, the panel provides visual cues that are not communicated through the corresponding text. Because a feature of graphic narratives is the visual presentation of facial expressions and situational context, graphic readers may have more information to generate emotion inferences than they gain from text narratives.

Few studies have investigated adult readers’ generation of emotion inferences during graphic narrative reading. Some prior research has investigated viewers’ understanding of characters’ emotions during visual narratives. For example, Clinton et al. (2018) examined how cinematography techniques (e.g., framing, point-of-view shots) influenced participants’ understanding of characters’ emotions during film. Future research could investigate how framing predicts a readers’ generation of emotion inferences during graphic narrative research (e.g., Ness-Maddox et al., 2022).

Figure 5.2.

Panel from Narrative B

Note: Image is from *Captain America* volume 3, issue 50 (Jurgens et al., 2002).

Predictions

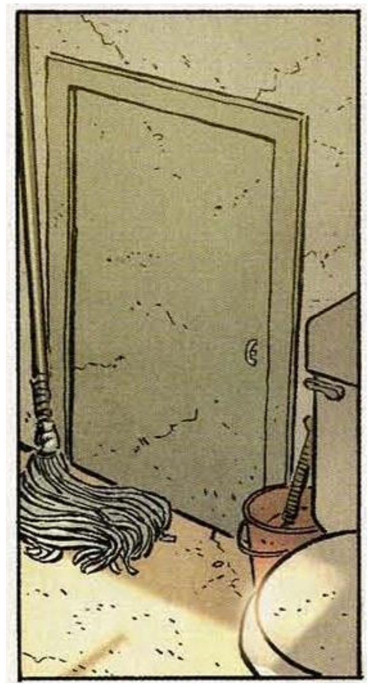
Post hoc analyses indicated that participants generated more predictions for text than non-linguistic graphic. Predictions involved participants' statements about what participants thought would happen next in the narratives. These results were qualified by a small effect of modality on predictions. While participants generated more backward-oriented or connecting inferences (i.e., anaphoric and bridging; Magliano et al., 1996) for non-linguistic graphic narratives, they generated more future-oriented inferences or predictions for text narratives.

Consider the panel in Figure 5.3. Panels prior to this panel show government agents looking around a building for a man after he appears to have escaped through a window in the restroom. After seeing the agents outside looking around the area, readers see this panel of another door in the restroom. The corresponding text of this panel says, "Inside the restroom was a small storeroom beside the toilet." Since participants generate access statements more for the non-linguistic graphic narratives than for text, a participant may view this panel and respond, "Here's a door with a mop and a bucket next to it, near a toilet." That is, they may simply describe the image. A participant reading the text, however, may make the predictive inference that the man running from the agents is about to emerge from behind this door.

Prior research has not directly compared participants generation of predictions for text and graphic narratives. Previous research does provide evidence that readers of graphic narratives do generate predictions based on their understanding of characters' narrative roles (i.e., agents and patients (Cohn et al., 2017; Cohn & Paczynski, 2013). Text features beyond characters' narrative roles may prompt readers to generate more predictions than they do for graphic, considering that text and non-linguistic graphic versions of the narratives in the current study include the same characters with the same narrative roles.

Figure 5.3.

Panel from Narrative C



Note: Image is from *Incredible Hulk* volume 2, issue 35 (Jones et al., 2002).

Goal Statements

Post hoc analyses also indicated that participants generated more goal statements for text narratives than non-linguistic graphic narratives. These results were qualified by a small effect of modality on goal statements. As with emotion, goals were not stated explicitly in the text versions, but actions and situational cues necessary to the plot were stated explicitly. Previously, Magliano et al. (2008) found that situational dimensions (i.e., time, place, characters, goals, events; Zwaan et al., 1995) influenced how participants created judgments about an imagined stranger's goals. Because actions and situational cues are stated more explicitly in text than graphic narratives, participants may be more easily prompted to infer characters' goals for text than non-linguistic graphic narratives (e.g., Magliano et al., 2008). It is worthwhile to note that

there was no difference in internal state inferences between text and non-linguistic graphic narratives. These results indicate that goal statements (e.g., "he decided to…"; "he wants to do…") are unique to inferences about other internal states (e.g., he recognizes him, she is thinking about…").

It is also important to note that text predictions and goals also have a significant positive correlational relationship (see external correlation matrix). Therefore, these two findings may be related. If a reader understands a character's goals, then the reader can more easily generate predictions of the character's future actions based on those goals . For example, Asiala et al. (2020) found that text readers used goal information to generate predictions about goal outcomes. One reason why readers may generate more predictions for text than non-linguistic graphic narratives is text narrative structure's emphasis on goals (e.g., Cohn, 2020; Johnson & Mandler, 1980; Rumelhart, 1977; Stein & Glenn, 1979). While text story grammars include the creation and attempt of goals, story grammars for visual narratives emphasize actions and events, without the inclusion of goals (e.g., Cohn, 2013; Cutting, 2016). Until the current study, no prior study has investigated readers' processing of narrative structure for text and graphic. While PINS Model (Cohn, 2020) emphasizes that narrative structure affects readers' online processing, participants did not say aloud many narrative elements, except goals and especially for text. Understanding characters' goals may guide readers to generate predictions of characters' actions aligned with those goals which is an open area of future research.

Affective Responses

Finally, participants generated more affective responses for text narratives than non-linguistic graphic narratives. Affective responses include statements about the participants' own emotional reaction to the narrative. These results were qualified by a small effect of modality on

affective responses. This result is somewhat surprising in light that participants generated more emotion inferences for non-linguistic graphic narrative than text narratives. However, while graphic narratives may provide the visual image of the character to provide enough information for the reader to generate the inference (e.g., Magliano et al., 2013), descriptions in text may encourage readers to think more about themselves and prompt affective responses. While this study was focused on narrative comprehension across modalities, future studies could investigate reader engagement across modalities.

Summary of Processes Generated during Narrative Comprehension

Overall, these findings support that comprehending non-linguistic graphic narratives is not easier, per se, than comprehending text (Coderre, 2020; Cohn & Magliano, 2020), but a reader's online processing may differ based on modality. During reading, graphic readers seem to process more surface form information (i.e., access visual information) which may in turn prompt readers to generate more emotion inferences about the visual information on the image (Ekman, 1992; Laubrock et al., 2018). Additionally graphic readers seem to generate more backward-oriented, connecting inferences to create a coherent representation of the narrative (Magaliano et al., 1996; Magliano et al., 2013; Trabasso et al., 1989; Trabasso & Magliano, 1996). Text readers, on the other hand, are more cognizant of characters' goals which may lead to more future-oriented predictions (Asiala et al.; 2020; Magliano et al., 2008; Magliano et al., 1996), as well as focus on how the information in the text may be causing them to feel for which they processed such information by generating affective responses (e.g., Koopman, 2015). Each of these findings is informative for understanding how readers process text and graphic narratives during reading. Next, I turn to discuss the results of testing within subjects differences across products participants produced during their recall responses after reading.

Differences in Products from Recall Responses

I hypothesized that participants would differ in reader comprehension products (i.e., text/image base and situation model) as measured by participants' recall across text and non-linguistic graphic narratives after controlling for participants' text print exposure and visual language fluency. I expected participants would include more accurate story information and more accurate description statements in their recall of non-linguistic graphic narratives than text narratives (Loughlin et al., 2015; Magliano et al., 2013). Additionally, I expected that participants would include more valid action inferences, valid emotion inferences, and valid general inferences in recall of non-linguistic graphic narratives than text narratives (Cohn, 2013; Cutting, 2016; Magliano et al., 2013). On the other hand, I expected that participants would include more valid goal inferences and valid internal state inferences in recall of text narratives than in recall of non-linguistic graphic narratives (Johnson & Mandler; 1980; Magliano et al., 2013; Rumelhart, 1977; Stein & Glenn,1979).

The results indicated that participants did differ in comprehension products across text and non-linguistic graphic narratives. However, the findings from post hoc analyses did not support my expectations. Unlike with the think-aloud responses in which participants generated more idea units for non-linguistic graphic narratives than text narratives, participants did not produce more idea units in their recall responses for one modality over the other. Next, I discuss the products in which participants differed across text and non-linguistic graphic narratives.

Story Information

Overall, participants included more accurate story information in their recall responses for text than in their recall responses for non-linguistic graphic narratives. Accurate story infor-

mation included information shown or stated explicitly in the narrative and did not require inferences. These results were qualified by a large effect of modality on recall of accurate story information. Recall of accurate story information is reflective of the participants' text base or image base representations (Kintsch, 1988; 1998). It is interesting that participants recalled more accurate story information for text narratives when they also generated more access statements for non-linguistic graphic narratives during the think-aloud task. This study did not include any direct research questions or hypotheses to examine how readers' cognitive processes impact their products. Thus, this interpretation is purely speculative and additional research would be warranted to understand this relationship further. Participants could also misinterpret the image (e.g., Cohn, 2021) and not incorporate accurate and necessary story information into their image base mental representation of the story, whereas with text narratives, the necessary narrative information is stated explicitly in the text which maps onto their text base mental representation of the text (Kintsch, 1988; 1998).

This finding may also be related to the finding that readers' generated more bridging inferences for non-linguistic graphic narratives than text narratives. For example, Magliano et al. (2017) found that generating bridging inferences of events not explicitly depicted distorts memory for explicit content. In the current study, participants generated more bridging inferences for the non-linguistic graphic narratives than the text narratives and then included more accurate story information for the text narratives than graphic narratives. However, again, the current study did not examine how readers' online processes relate to their offline comprehension products and cannot make causal conclusions. Additional research could investigate how readers' online processing for text and non-linguistic graphic narrative may lead to their offline comprehension products.

Alternatively, participants may have recalled more accurate story information for text than non-linguistic graphic narratives because the text narratives explicitly communicated story information. For example, consider the panel from Narrative D and its corresponding text sentence in Figure 5.4. In the text version, the sentence explicitly states that Loki gestured for Thor to follow him. In the non-linguistic graphic version, the viewer must have prior knowledge for what Loki's hand gesture means and infer what intentions Loki is trying to communicate to Thor (Cohn, 2016). In the text, the reader is provided specific information about Loki's intentions without additional prior knowledge or inferences needed. Thus, participants may have remembered more accurate story information for text because the text versions of the narratives more explicitly and more specifically communicated story information than the non-linguistic graphic versions. That is, participants may have remembered fewer details of story information for non-linguistic graphic narratives because those details were not specifically communicated to them through the images.

Figure 5.4

Comparison of Specificity in Non-Linguistic Panel and Corresponding Text Sentence

Non-Linguistic Graphic Version

Text Version

Thor looked at his sword, frowning, and Loki gestured for Thor to follow him.

Note: Image is from *Thor* volume 2, issue 44 (Jurgens et al., 2002).

It is important to note that participants did not differ in their inclusion of inaccurate story information between the modalities during their recalls. These findings indicate that participants were not recalling more false or inaccurate information about the non-linguistic graphic narratives compared to text, but that they were not recalling as many accurate details about the story as they were for text narratives.

This finding suggests that readers may have differences between their text base and image base representations for text and graphic narratives as well as content differences in their situation model. Prior studies support that children construct both the image base and situation model for graphic narratives (Meringoff, 1980; Poulson et al., 1979). Results of the current study somewhat align with Meringoff's study, which found that when children view a narrative with more linguistic information (i.e., picture stories with text), they recall more story information. Similarly, results of the current study indicate that adult participants included more accurate story information for recall of text narratives than non-linguistic graphic narratives. Meringoff also found that when children view a narrative with more visual information (i.e., animated film with linguistic audio), they recall more actions. In the current study, adult participants did not include more action inferences in their recall of one modality over the other. Future research could investigate whether children may be more influenced by modality and exhibit more or other differences in their comprehension products across text and non-linguistic graphic narratives than the adult readers did in the current study.

Emotion and Internal State Inferences

Post hoc analyses indicated that participants included more emotion inferences in recall responses for non-linguistic graphic narratives than text narratives. These results indicated that modality had a small effect on emotion inferences. This finding is not surprising considering that

participants generated more emotional inferences for non-linguistic graphic narratives than text narratives during the think-aloud task. This finding suggests that participants included the emotional inferences they generated during online processing of the narrative into their mental representation of the narrative. However, as mentioned previously, this study did not include any direct research questions or hypotheses to examine how readers' cognitive processes impact their products. Thus, this interpretation is purely speculative and additional research would be warranted to understand this relationship further. Nonetheless, when prompted to recall the narrative, participant could have included the emotion inferences they incorporated during reading into their mental representation (Graesser et al., 1994; Kintsch, 1998; Magliano et al., 2013). These results did not indicate other differences among inferences included in recall responses for text and non-linguistic graphic narratives, only for emotion inferences. This finding could suggest that modality affects participants' recall of emotional information inferred during reading, and that graphic narratives presentation of visual context through facial expressions and setting leads to this difference (Ekman, 1992; Gnepp, 1983; Laubrock et al., 2018; Ness-Maddox et al., 2022).

It is important to note that while a test of significance cannot be conducted to determine mean differences, participants did not include any invalid internal state inferences in their recall responses for text narratives but did include invalid internal state inferences in their recall responses for non-linguistic graphic narratives. Previously, Lewis et al. (1994) found that linguistic information can guide children to generate internal state inferences while viewing picture stories. The current study involved adult participants, not children, who showed no differences in the online generation of internal state inferences between modalities, but only that participants who included invalid internal state inferences within their offline recall of non-linguistic graphic nar-

ratives. These findings may suggest that adult readers of graphic narratives are more likely to incorrectly infer characters' internal or mental states from images than they are for text narratives and integrate these invalid inferences into their mental representations. However, this explanation would need to be further tested in future research. These results also did not indicate any other differences among invalid inferences.

Metacognition

Finally, participants included more metacognitive comments in their recall responses for text narratives than non-linguistic graphic narratives. These results were qualified by a small effect of modality on metacognition. It is worthwhile to note that metacognitive comments did include statements about the participants lack of understanding or memory of the story, they also included any statements about the participants' understanding of the narrative. Therefore, metacognitive comments do not necessarily indicate confusion or lack of understanding, only that the participant discussed their own understanding (.van den Broek et al., 2001). It would be inaccurate to conclude that differences in metacognitive comments indicates comprehension or lack thereof. Thus, these results only indicate that participants verbalized their awareness of their own metacognition while recalling text narratives more than recalling non-linguistic graphic narratives. However, additional research could include further investigation into readers' awareness of their understanding of text compared to graphic narrative memory.

Overall Conclusion of Findings

Overall, participants' comprehension processing during think-aloud tasks and products produced during recall tasks were found to be more similar across text and non-linguistic graphic narratives than they are different. However, the differences revealed by results of this study seem to stem from features of the modalities (Magliano et al., 2018). The panels in the non-linguistic

graphic versions of the narratives present more information through images than what is conveyed in the text versions of the narratives. This additional information may include facial expressions, setting, and potentially distracting objects and characters (e.g., Gnepp, 1983; Laubrock et al., 2018). Readers must interpret the images and discern what information is necessary to comprehending the story. On the other hand, the text narratives used in this study included necessary information explicitly stated in the text. Thus, inferences were still necessary to generate in order to connect the information across sentences; however, these texts may not have required an additional need to generate other inferences needed to fill implicit gaps in the text. This may have then have led readers to interpret and discriminate necessary information less for text than for non-linguistic graphic narratives. These texts may have also provided the readers the need to identify the characters' goals which may allow them to think forward in the story and generate predictions about their goals and corresponding outcomes (Magliano et al., 1996). These features of text and non-linguistic graphic narratives may have led to the differences found in online processing and offline products.

Online Processing Differences

Overall, for the non-linguistic graphic narratives used in this study, participants seem to have generated inferences in order to connect the sequence of panels into a coherent story, using backward-oriented inferences (i.e., anaphoric and bridging; e.g., Magliano et al., 1996). For text narratives, participants appeared more cognizant of characters' goals which may have also encouraged them to generate future-oriented inferences (i.e., predictions). For non-linguistic graphic narratives, participants may also be attending to information beyond what is included in the text versions, such as facial expressions and setting, which can then prompt participants to generate more emotional inferences (Ekman, 1992; Laubrock et al., 2018).

Offline Product Differences

Participants included more accurate information in their offline mental representations of text narratives than non-linguistic graphic narratives. As mentioned above (see *Access*), perhaps participants were able to attend to and incorporate necessary story information into their mental representation for text narratives than for non-linguistic graphic narratives because the amount of information is less in the text versions of these narratives than the non-linguistic graphic narrative. Therefore, they included explicit content and story information more in their mental representations of the text narratives.

The only difference in inference inclusion during recall was emotion inferences. This difference in offline products may stem from differences in modality features between text and graphic narratives (Ekman, 1992; Laubrock et al., 2018). For instance, the non-linguistic graphic versions of the narratives present additional context information not included in the text versions. While this additional information may inhibit readers from attending to and incorporating important information into their mental representation, it does allow readers to develop understandings of characters' emotional states during and after reading.

Thus, overall, the findings for this study support the modality of narratives on reader comprehension processing and products. These findings have implications for theories of narrative comprehension.

Implications for Comprehension Across Text and Non-Linguistic Graphic Narrative Modalities

CI Model

The results of this study provide additional evidence that readers of both text and non-linguistic graphic narratives construct text/image base and situation model mental representations

as the CI Model suggests (Kintsch, 1988; 1998; Kintsch & Van Dijk, 1978). However, those mental representations may look different depending on the modality. For example, findings from the current study support participants' construction of a richer text base representations by including more accurate story information in their recall responses. These findings may suggest that readers integrate explicit linguistic information more efficiently from text into their text base mental representations than they do visual information from graphic narratives into their image base mental representation. Graphic narratives may also depict a character or an event (i.e., explicitly show a character or event), but a reader may need to interpret the visual information as that character or event. Therefore, constructing the image base representation may require more cognitive processing than constructing the text base representation.

Overall, participants did not differ greatly across inferences included in their recall responses, and, for the most part, their situation model representations were similar across text and non-linguistic graphic narratives. Participants did, however, include more emotional inferences in their recalls for non-linguistic graphic narratives than text. Participants also generated more emotional inferences during the think-aloud task. That is, during online processing, participants were more cognizant of characters' emotions for non-linguistic graphic narratives than text narratives. Readers generated emotional inferences during online processing, and those inferences may have remained salient in their offline comprehension products, or mental representations post-reading. Readers of graphic narratives, thus, may have developed their situation model representations with more of an understanding of characters' emotions than text readers. Additional research, however, is needed to determine whether the emotion inferences readers generate during online processing are related to or predict readers' inclusion of emotion inferences in their offline products of the narrative.

Additionally, the CI Model states that mental representations are constructed by features of the text activating the readers' background knowledge (Kintsch, 1998; Kintsch & Van Dijk, 1978; McNamara & Magliano 2009). Readers' background knowledge of emotions may have been more activated by visual information in the graphic narratives used in this study than the information provided in the text narratives. That is, the graphic narratives may have provided visual information in the form of facial expressions and situational context, and this information may have activated readers' socioemotional knowledge. Therefore, when considering how a reader constructs their mental representations of a narrative, it may be worthwhile to consider how features of the modality provide different opportunities for activation of a reader's knowledge and how well the reader can then integrate that knowledge and information from the narrative into their mental representations. Implications of these findings within the front-end/back-end framework (Magliano et al.., 2013) will help explain how modality features influence front-end processes which in turn may influence the activation of knowledge.

Front-end/Back-end Framework

Magliano et al. (2013) state that while readers engage in different front-end processes for different narratives depending on the information presented (i.e., text and linguistic or visual), readers still generate the same back-end processes regardless of the modality. That is, readers may take in different types of information when reading graphic and text narratives, but readers process or think about that information similarly regardless of the type of information. These results of this study support that readers do generate the same back-end processes for text and non-linguistic graphic narratives, but not always to the same extent across the modalities. For example, in the current study, readers generated bridging inferences for both text and non-linguistic

graphic narratives, but they generated more bridging inferences (as well as anaphoric and emotion inferences) for non-linguistic graphic narratives than text.

Readers of different narrative modalities may generate back-end inferences to different extents for two related reasons. First, features of graphic narratives prompt readers to generate more connecting or backward-looking inferences than for text. As discussed, graphic narratives are composed of sequential images which the reader must connect to develop a coherent mental representation of the narrative depicted in the images. Thus, readers generate more bridging and anaphoric inferences for non-linguistic graphic narratives than text narratives. Second, engaging in scene gist processing and object/character processing (i.e., front-end processes specific to visual narratives; Magliano et al., 2013) may specifically activate socioemotional knowledge in the readers' background knowledge and prompt the generation of emotional inferences while reading. Thus, it is accurate to say that readers do generate back-end processes regardless of the narrative modality, but front-end processes specific to a modality may lead to the activation of different knowledge and the generation of various types back-end processes to different extents.

iLC Framework

The iLC Framework (Kendeou et al., 2020) is based on prior research supporting that readers have a general inference ability that spans across narrative modalities (e.g., Kendeou et al., 2008; Kendeou & van den Broek, 2005). Kendeou and colleagues list four specific types of inferences readers generate for both text and graphic modalities: anaphoric, bridging, elaboration, and emotion inferences. While participants in this study did generate anaphoric, bridging, elaboration, and emotion inferences for both text and non-linguistic graphic narratives, they generated more anaphoric, bridging, and emotion inferences for non-linguistic graphic narratives

than text. Participants did not differ in elaborations across narrative modalities. Additionally, results of this study indicated significant positive correlational relationships for text and non-linguistic graphic anaphoric, bridging, and elaboration inferences (see external spreadsheet). Notably, emotion inferences for text and non-linguistic graphic narratives were not related. While readers may have a general inference ability that extends across modalities, the extent to which readers generate emotion inferences may be less affected by this general inference ability and more affected by features of graphic narratives, and therefore, modality differences.

Internal state inferences were included in the online processing coding scheme for this study after pilot coding indicated that participants generated inferences about the characters' internal (i.e., mental) state that did not relate to the characters' emotions or goals. That is, participants generated inferences about what characters were thinking. Internal state inferences for text and non-linguistic graphic narratives also had a significant, positive relationship. Internal state inferences may be a useful category to consider in addition to anaphoric, bridging, elaborations, and emotion inferences already included in the iLC Framework (Kendeou et al., 2020). Overall, these results provided additional evidence for general inference ability across text and graphic modalities in that anaphoric, bridging, elaborations, and internal state inferences were all significantly related across modalities. However, while readers may have a general inference ability that extends across modality, readers may still differ in the extent they generate these inferences across text and graphic narratives.

PINS Model

The PINS Model (Cohn, 2020) emphasizes the importance of simultaneously processing semantic content of the narrative along with narrative structure. Readers access semantic and structural information from the narrative and based on this semantic and structural information,

readers generate predictions. As they continue to access information that confirms or contradicts their predictions, readers update their mental representations. The results of this study support this notion for how readers access both text and graphic semantic and narrative structure information, generate predictions, and update their understanding of the narratives as they read (Cohn, 2020). Participants produced more access statements during their think-aloud responses for non-linguistic graphic narratives than text and did not differ on narrative statements (i.e., events, characters, scene) across modalities. Participants in this study generated more predictions for text than non-linguistic graphic narrative. Notably, predictions for text and non-linguistic graphic narratives were also not related, and participants did not differ in the extent they generated update responses during the think-aloud task. Therefore, a reader may have a tendency to generate predictions while reading text narratives that is unrelated to their tendency for generating predictions for graphic narratives. Results indicated that updates were related across text and non-linguistic graphic narratives. Therefore, while readers may differ in their generation of predictions across text and graphic narratives, they update their mental representations similarly across modalities while reading.

Narrative structure information may also be easier for readers to access with text narratives and then consequently to generate more predictions than with graphic narratives. While participants in this study did generate statements about the structure of narratives during think-aloud tasks, they generated more goal statements than statements about events, scene, or characters. Additionally, participants generated more goal statements for text than non-linguistic graphic narratives, and these goal statements were not related across text and non-linguistic graphic narratives. Text narrative structure tends to focus on the creation and attempts of goals while narrative structure for visual narratives tend to focus on actions and events (e.g., Johnson & Mandler,

1980; Rumelhart, 1977; Stein & Glenn, 1979). Previous studies have not compared how readers process narrative structure while reading text and non-linguistic graphic narrative. The current study did compare statements about narrative structure across text and non-linguistic graphic narratives, but results did not indicate other significant differences statements about narrative structure other than goals. Statements about events and characters were positively related across modalities, but statements about narrative structure across modalities were fairly sparse compared to other processes observed during think-aloud tasks. Readers may access information about structure from both text and non-linguistic graphic narratives; however, this structural information may not be not as salient as goals in narrative which participants verbalized more in their think-aloud responses, especially for text.

Limitations

Although the findings from this study supported and extended previous research, it is also not without its limitations. First, with this study, I included only adults readers as my participants. The purpose of investigating comprehension of adult readers was to increase the likelihood that participants would not have difficulty comprehending the text and non-linguistic graphic narratives. Therefore, results of this study do not necessarily extend to younger readers.

Additionally, this study is limited by the narrative materials used specifically for this study. While I choose these non-linguistic graphic materials because they are graphic narratives a reader may actually read, the text versions are less naturalistic. The non-linguistic graphic versions were created for audience consumption and published. The text versions, however, are not typical of text narratives. For example, most text narratives do describe characters' emotions rather than only describing facial expressions. For purposes of this study to compare online pro-

cessing and offline products, it was necessary for the text versions to not explicitly provide information that readers of the non-linguistic graphic versions would need to infer. However, these findings may not accurately and completely reflect the differences in online processing and offline products due to modality that occur with readers of narratives outside this study.

Finally, this study investigated different types of inferences as back-end processes but did not investigate the other types of process. For example, Magliano et al. (2013) identify the a back-end process of event-segmentation. Previous research has investigated event segmentation during graphic narrative reading (e.g., Kaiser, 2020) and between visual narratives and text narratives (e.g., Zacks et al., 2009). In light of the findings from the current study that identified participants generated more backward-oriented inferences (i.e., anaphoric and bridging inferences, possibly to create coherence across panels, future research may investigate how readers segment text and non-linguistic graphic narratives into events and how this relates to readers' coherent mental representations as they read text and non-linguistic graphic narratives.

Finally, the current study is limited in the scope of its conclusions. Because I did not investigate or analyze the relationship between readers' online processes with their offline products, I cannot conclude whether readers' processes generated during reading influenced the nature of their text/image base or situation model representations. The relationship between online processes and offline products is supported by the CI Model and front-end/back-end framework (Kintsch, 1988; 1998; Magliano et al, 2013); however, future empirical research is necessary to investigate the nature of possible relationships between readers' online processes and offline products further.

Implications for Research and Future Studies

This study extends prior research by directly comparing non-linguistic graphic narratives (i.e., only visual information) with text narratives (i.e., linguistic information). No prior study has directly compared comprehension processes and products of visual and linguistic modalities to the extent as the current study has. The findings extend the research in the field of narrative comprehension by providing evidence that while overall readers process and remember text and non-linguistic graphic narratives similarly, modality does influence the extent to which readers access narrative information, generate coherence building inferences, generate emotion inferences, understand characters' goals, and generate predictions. Additionally, the findings suggest that text readers have more developed text base representations than image base representations, and graphic readers include more emotion inferences in their situation models than text readers. The field of graphic narrative comprehension research is quickly growing, and there are many opportunities to extend the research findings here. I will highlight three areas of opportunities: exploring features of narrative modalities, replication with younger readers and diverse populations, and investigation of reader comprehension processing profiles.

Features of Narrative Modalities

The current study supports that modality can influence a reader's online processing and offline products. Future research could extend the findings of the current study by investigating the features of text and non-linguistic graphic modalities that may contribute to these comprehension processing and product differences. For example, future research could investigate how readers integrate situational dimensions (i.e., setting, time, events, characters; Zwaan et al., 1995) from non-linguistic graphic and text narratives into their situation model representation and whether these situational dimensions explain the difference in goal statements and consequently

predictions for text narratives over non-linguistic graphic narratives (e.g., Asiala et al., 2020; Magliano et al., 2008).

Likewise, further research can investigate how framing and focus on characters' facial expressions predicts a readers' generation of emotion inferences during graphic narrative research (e.g., Ness-Maddox et al., 2022). For example, Ness-Maddox et al. (2022) presented preliminary findings that framing of graphic narrative panels (e.g., long shot, full shot, close up shot) predicted whether readers of non-linguistic graphic narratives generated an emotional inference. Preliminary findings for this presentation support that graphic panels that closely frame characters' faces (i.e., medium close up, close up, extreme close up shots) prompt readers to generate an emotional inferences.

Replication with Younger Readers and Diverse Groups

This study could be replicated with children. In this study, I measured participants' exposure to text print through the ART (Acheson et al., 2008) and their background experience with graphic narratives through the VLF index (Cohn, 2014). While the ART measure may not be suitable for children, there are other measures which could be used to evaluate children's experience reading text (e.g., Woodcock-Johnson Reading Fluency subtest; Woodcock et al., 2001; 2014). The VLF index, to my knowledge, has only been used as a measure for adult participants. Studies with children may need to adapt and validate a child version of the VLF index. A future study replicating these results would be informative for the iLC Framework (Kendeou et al., 2020) and related interventions based on improving children's inference generation for text.

Additionally, much of the research discussed here investigated comprehension of typically developing (TD) readers. In this study, I did not measure participants' language ability and socioemotional skills, beyond screening for decoding skills. Individuals diagnosed with Autism

Spectrum Disorder and language deficits have difficulty comprehending non-linguistic graphic narratives as they do text narratives (e.g., Coderre, 2020) and have difficulty interpreting facial expressions and social situations (e.g., Zhang et al., 2022). Perhaps this population would not exhibit differences in text and non-linguistic graphic narratives regarding the generation of emotional inferences. Future research could compare emotional inferences across text and non-linguistic graphic narratives for both TD and ASD samples.

Comprehension Processing Profiles for Graphic Reading

Finally, prior research for text has found reader profiles based on their online processing tendency during think-aloud tasks (e.g., Karlsson et al., 2018). That is, both younger readers and college students tend to generate causal inferences, generate elaborations, or paraphrase the text while reading (Carlson et al., 2014; Davison et al., 2018; McMaster et al., 2012). Future directions of the research conducted here could investigate whether readers fall into similar reader profiles for text and for non-linguistic graphic narratives. That is, if a reader tends to paraphrase while reading text, does that same reader also tend to describe while reading graphic narratives? Or could a reader have one processing profile for text and another profile for graphic narratives? This research, especially conducted with children, would also be informative for the iLC Framework (Kendeou et al., 2020) and related interventions. Understanding how a reader may tend to process text and graphic narratives would influence how they would respond to interventions to scaffold inference skills. Additionally, future research could investigate whether the online processes readers generate during reading (e.g., bridging inferences, emotion inferences) influence the type of content readers include in their text/image base and situation model representations (e.g., accurate story information, valid emotion inferences).

Implications for Practice

This study also has implications for practice. The iLC Framework (Kendeou et al., 2020) centers on the idea that readers have a general inference ability that spans across narrative modality (Kendeou et al., 2008; Kendeou & van den Broek, 2005). Based on this general inference ability, graphic narratives could be used in interventions to scaffold children's inference generation in text narratives. Overall, results of the current study support the idea of general inference ability across modalities but also supports that modality may influence the type of inference a reader may generate. Rather than use graphic narratives to scaffold inference generation in general, perhaps graphic narratives could specifically be used in interventions to build readers' ability to generate inferences that connect ideas across the text (e.g., bridging, anaphoric; Brenna, 2012). Graphic narratives may also be used to scaffold social and emotional intelligence by highlighting visual and situational cues that prompt emotion inferences. At the same time, text may also be used to scaffold understanding of characters' goals and predicting events and behavior, which may be a beneficial skill outside the context of narrative reading and in real life situations. Thus, future intervention studies are needed to determine whether text and graphic narrative modalities could each improve young readers generation of different inferences.

Conclusions

This study provides support that readers generate online processes to different extents for text and graphic narratives. Additionally, readers' offline mental representations of the narratives differ in the content included across text and graphic narratives. These findings provide additional evidence that the Visual Ease Assumption is unsupported (e.g., Coderre, 2020). That is, this study provides no evidence that comprehending graphic narratives is easier than comprehending text narratives. However, while comprehension of one modality may not be easier, per

se, than the other does not mean differences cannot occur. Broadly, graphic readers attend to visual information from panels and connect panels into a coherent narrative through bridging and anaphoric inferences. This difference in the generation of anaphoric and bridging inferences across modalities may indicate that readers integrate information into their text base representation for text but need to generate inferences to integrate the same information into their situation model representation for graphic narratives. For text narratives, readers focused on motivations or goals of characters which may prompt predictions. For graphic narratives, readers focused on emotions. Consequently, emotion is more salient in graphic readers' offline products as well. Text readers, however, have richer text base representations of story details from the narrative than graphic readers. Thus, this study supports that modality has an effect on readers' mental representations of the narrative. Visual and linguistic information presented in graphic and text narratives influence how readers process a narrative and what they include in their offline representations of the narrative. In conclusion, the form of information (i.e., linguistic, visual) readers take in while reading influence how we think about and remember that information.

REFERENCES

Abdi (2010). The Greenhouse-Geisser correction. In Neil Salkind (Ed.), *Encyclopedia of Research Design.* Sage.

Acheson, D. J., Wells, J. B., & MacDonald, M. C. (2008). New and updated tests of print exposure and reading abilities in college students. *Behavior Research Methods, 40*, 278-289.

Asiala, L. K. E., Chan, G. C., Kurby, C. A., & Magliano, J. P. (2020). The role of goals and goal barriers in predicting the outcomes of intentional actions in the contexts of narrative text. *Journal of Cognitive Psychology, 32*(1), 82–92. https://doi.org/10.1080/20445911.2019.1690494

Botas, G. (2017). Differences in strategy use in the reading comprehension of narrative and science texts among students with and without learning disabilities. *Learning Disabilities: A Contemporary Journal*, *15*(1), 139-162.

Brenna, B. (2012). How graphic novels support reading comprehension strategy development in children. *Literacy, 47*(2), 88-94.

Carlson, S., Seipel, B., & McMaster, K. (2014). Development of a new reading comprehension assessment: Identifying comprehension differences among readers. *Learning And Individual Differences, 32*, 40-53. doi: 10.1016/j.lindif.2014.03.003

Clinton, J. A., Briner, S. W., Sherrill, A. M., Ackerman, T., & Magliano, J. P. (2018). The role of cinematic techniques in understanding character affect. *Scientific Study of Literature*, *7*(2), 177–202. https://doi.org/10.1075/ssol.16019.cli

Clinton, V., Taylor, T., Bajpayee, S., Davison, M. L., Carlson, S. E., & Seipel, B. (2020).

Inferential comprehension differences between narrative and expository texts: A systematic review and meta-analysis. *Reading and Writing: An Interdisciplinary Journal.* https://doi.org/10.1007/s11145-020-10044-2

Coderre, E.L. (2020). Dismantling the "visual ease assumption:" A review of visual narrative processing in clinical populations. *Topics in Cognitive Science, 12*, 224-255.

Coderre, E.L., Cohn, N., Slipher, S.K., Chernenok, M., Ledoux, K., and Gordon, B. (2018). Visual and linguistic narrative comprehension in autism spectrum disorders: Neural evidence for modality-independent impairments. *Brain and Language, 186*, 44-59.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Cohen, J. (1992). A Power Primer. *Psychological Bulletin, 112*, 155-159. https://doi.org/10.1037/0033-2909.112.1.155

Cohn, N. & Kutas, M. (2017). What's your neural function, visual narrative conjunction? Grammar, meaning, and fluency in sequential image processing. *Cognitive Research: Principles And Implications, 2*(1), 1-13. doi: 10.1186/s41235-017-0064-5

Cohn, N. & Maher, S. (2015). The notion of the motion: The neurocognition of motion lines in visual narratives. *Brain Research, 1601*, pp.73-84.

Cohn, N. (2013). Visual narrative structure. *Cognitive Science, 34*, 413-452.

Cohn, N. (2014). The Visual Language Fluency Index: A measure of "comic reading expertise." Visual Language Lab: Resources. www.visuallanguagelab.com/resources.html. First posted online April 2014.

Cohn, N. (2016). *The Visual Narrative Reader.* Bloomsbury.

Cohn, N. (2020). Your brain on comics: A cognitive model of visual narrative comprehension.

Topics in Cognitive Science, 12, 352-386.

Cohn, N. (2021). A starring role for inference in the neurocognition of visual narratives. *Cogn. Research, 6*(8).

Cohn, N. (2021). *Who understands comics?.* New York, NY: Bloomsbury Academic.

Cohn, N. and Foulsham, T. (2020). Zooming in on the cognitive neuroscience of visual narrative. *Brain and Cognition, 146.*

Cohn, N. and Magliano, J.P. (2020) Editors' introduction and review: Visual narrative research: An emerging field in cognitive science. *Topics in Cognitive Science, 12*, 197-223.

Cohn, N. and Paczynski, M. (2013). Prediction, events, and the advantage of Agents: The processing of semantic roles in visual narrative. *Cognitive Psychology, 67*(3), 73-97.

Cohn, N., & Wittenberg, E. (2015). Action starring narratives and events: Structure and inference in visual narrative comprehension. *Journal Of Cognitive Psychology, 27*(7), 812-828. https://doi.org/10.1080/20445911.2015.1051535

Cohn, N., Jackendoff, R., Holcomb, P., and Kuperberg, G. (2014). The grammar of visual narratives: Neural evidence for constituent structure in visual narrative comprehension. *Neuropsychologia. 64*, 63-70.

Cohn, N., Paczynski, M., and Kutas, M. (2017). Not so secret agents: Event-related potentials to semantic roles in visual event comprehension. *Brain and Cognition, 119*, 1-9.

Cook, M. P. (2017). Now I "see": The impact of graphic novels on reading comprehension in high school English classrooms. *Literacy Research and Instruction, 56*(1), 21-53.

Coopmans, C.W. and Cohn, N. (2022). An electrophysiological investigation of co-referential processes in visual narrative comprehension. *Neuropsychologia, 172.*

Cutting, J. (2016). Narrative theory and the dynamics of popular movies. *Psychonomic Bulletin*

& Review, 23(6), 1713-1743. https://doi.org/10.3758/s13423-016-1051-4

Davison, M. L., Biancarosa, G., Carlson, S. E., Seipel, B., & Liu, B. (2018). Preliminary findings on the computer-administered Multiple-choice Online Causal Comprehension Assessment, a diagnostic reading comprehension test. *Assessment for Effective Intervention, 43*(3), 169-181.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion, 6*, 169–200.

Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2009). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.

Foulsham, T. and Cohn, N. (2021). Zooming in on visual narrative comprehension. *Memory & Cognition, 49*, 451-466.

Foulsham, T., Wybrow, D., and Cohn, N. (2016). Reading without words: Eye movements in the comprehension of comic strips. *Applied Cognitive Psychology, 30*, 566-579.

Gnepp, J. (1983). Children's social sensitivity: Inferring emotions from conflicting cues. *Developmental Psychology, 19*(6), 805-814.

Graesser, A., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101*(3), 371-395. https://doi.org/10.1037/0033-295x.101.3.371

Hellmann, E. (2018). Uncovering what readers know: Understanding readers' online and offline processes for identifying story elements [ProQuest Information & Learning]. In *Dissertation Abstracts International Section A: Humanities and Social Sciences* (Vol. 79, Issue 4–A(E)).

Hoover, W.A. & Gough, P.B. (1990). The simple view of reading. *Reading and Writing: An*

Interdisciplinary Journal, 2, 127-160.

Huck, S., 2012. *Reading statistics and research.* 6th ed. Boston, MA: Pearson Education.

Huff, M., Rosenfelder, D., Oberbeck, M., Merkt, M., Papenmeier, P., Meitz, T.G.K. (2020). Cross-codal integration of bridging-event information in narrative understanding. *Memory & Cognition, 48*, 942-956.

Hutson, J.P., Magliano, J.P., and Loschky, L.C. (2018). Understanding moment-to-moment processing of visual narratives. *Cognitive Science, 42*, 2999-3033.

Johnson, N.S., and Mandler, J. M. (1980). A tale of two structures: Underlying and surface forms in stories. *Poetics, 9*, 51-86.

Jones, B., Romita, Jr., J., Palmer, T., & Alonso, A. (2002, February). “Silent Running.” *The Incredible Hulk, 2*(35). [Comic book]. Marvel Comics.

Jurgens, D, Immonen, S., Koblish, S., Kemp, D., Starkings, R., Brevoort, T. (2002, February). “As Lives a Memory…” *Thor, 2*(44). [Comic book]. Marvel Comics.

Jurgens, D., Layton, B., & Klein, T. (2002, February). “Silent Night-Silent Morning.” *Captain America, 3*(50). [Comic book]. Marvel Comics.

Kaiser, E. (2020). Linguistic consequences of event segmentation in visual narratives: Implications for prominence. *Language, Cognition and Neuroscience, 35*(3), 402–408. https://doi.org/10.1080/23273798.2019.1667000

Karlsson, J., van den Broek, P., Helder, A., Hickendorff, M., Koornneef, A., & van Leijenhorst, L. (2018). Profiles of young readers: Evidence from thinking aloud while reading narrative and expository texts. *Learning and Individual Differences, 67*, 105-116.

Kendeou, P., & van den Broek, P. (2005). The Effects of Readers’ Misconceptions on

Comprehension of Scientific Text. *Journal of Educational Psychology*, *97*(2), 235–245. https://doi.org/10.1037/0022-0663.97.2.235

Kendeou, P., Bohn-Gettler, C., White, M.J., and van den Broek, P. (2008). Children's inference generation across different media. *Journal of Research in Reading*, *31*(3), 259-272.

Kendeou, P., McMaster, K.L., Butterfuss, R., Kim, J., Bresina, B., and Wagner, K. (2020). The inferential language comprehension (iLC) framework: Supporting children's comprehension of visual narratives. *Topics in Cognitive Science*, *12*, 256-273.

Kincaid, J.P., Fishburne, R.P., Rogers, R.L., & Chissom, B.S. (1975). Derivation of new readability formulas (automated readability index, fog count, and flesch reading ease formula) for Navy enlisted personnel. Research Branch Report 8–75. Chief of Naval Technical Training: Naval Air Station Memphis.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A constructive integration model. *Psychological Review*, *95*, 163–182.

Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*(5), 363.

Kintsch. W. (1998). *Comprehension: A paradigm for cognition.* Cambridge, MA: Cambridge University Press.

Kline, R. B. (2011). *Principles and practice of structural equation modeling*, 4th ed. Guilford Press.

Koopman, E. M. (Emy). (2015). How texts about suffering trigger reflection: Genre, personal factors, and affective responses. *Psychology of Aesthetics, Creativity, and the Arts*, *9*(4), 430–441. https://doi.org/10.1037/aca0000006

Kraal, A., Koornneef, A., Saab, N., & van den Broek, P. (2018). Processing of expository and

narrative texts by low- and high-comprehending children. *Reading And Writing, 31*(9), 2017-2040. doi: 10.1007/s11145-017-9789-2

Laubrock, J., Hohenstein, S., & Kümmerer, M. (2018). Attention to comics:Cognitive processing during the reading of graphic literature. In A. Dunst,J. Laubrock, & J. Wildfeuer (Eds.), *Empirical comics research: Digital, multi-modal, and cognitive methods* (pp. 239–263). Routledge.

Levoy, B. (2003). Audio textbooks: An important tool for those with learning-related & low vision problems. *Journal of Behavioral Optometry, 14*(5), 121–122.

Lewis, C., Freeman, N.H., Hagestadt, N.H., and Douglas, H. (1994). Narrative access and production in preschoolers' false belief reasoning. *Cognitive Development, 9,* 397-424.

Loughlin, S., Grossnickle, E., Dinsmore, D., and Alexander, P. (2015). “Reading” paintings: Evidence for trans-symbolic and symbol-specific comprehension processes. *Cognition and Instruction, 33*(3), 257-293.

Magliano, J. P., & Millis, K. K. (2003). Assessing reading skill with a think-aloud procedure and latent semantic analysis. *Cognition and Instruction, 21*, 251–283.

Magliano, J. P., Dijkstra, K., & Zwaan, R. A. (1996). Generating predictive inferences while viewing a movie. *Discourse Processes, 22*(3), 199–224. https://doi.org/10.1080/01638539609544973

Magliano, J. P., Kopp, K., McNerney, M. W., Radvansky, G. A., & Zacks, J. M. (2012). Aging and perceived event structure as a function of modality. *Aging, Neuropsychology, and Cognition, 19*(1–2), 264– 282. https://doi.org/10.1080/13825585.2011.633159.

Magliano, J. P., Larson, A. M., Higgs, K., & Loschky, L. C. (2016). The relative roles of

visuospatial and linguistic working memory systems in generating inferences during visual narrative comprehension. *Memory & Cognition, 44*, 207–219.

Magliano, J. P., Skowronski, J. J., Britt, M. A., Güss, C. D., & Forsythe, C. (2008). What do you want? How perceivers use cues to make goal inferences about others. *Cognition, 106*(2), 594–632. https://doi.org/10.1016/j.cognition.2007.03.010

Magliano, J., Kopp, K., Higgs, K. & Rapp, D. (2017). Filling in the gaps: Memory implications for inferring missing content in graphic narratives. *Discourse Processes, 54*(8), 569-582.

Magliano, J.P. and Graesser, A.C. (1991). A three-pronged method for studying inference generation in literary text. *Poetics, 20*, 193-232.

Magliano, J.P., Loschky, L.C., Clinton, J.A., & Larson, A.M. (2013). Is reading the same as viewing? An exploration of the similarities and differences between processing text- and visually based narratives. In B. Miller, L. Cutting, and P. McCardle (Eds.), *Unraveling the Behavioral, Neurobiological, & Genetic Components of Reading Comprehension,* (pp. 78-90). Brookes Publishing Co.

Magliano, J.P., McCrudden, M.T., Rouet, J., and Sabatini, J. (2018). The modern reader: Should changes to how we read affect research and theory? In Schober, M.F., Rapp, D.N., and Britt, M.A. (Eds.) (2018). *The Routledge Handbook of Discourse Processes* (2nd Ed.). Taylor & Francis.

Manfredi, M., Cohn, N., and Kutas, M. (2017). When a hit sounds like a kiss: An electrophysiological exploration of semantic processing in visual narrative. *Brain & Language, 169*, 28-38.

McCloud, S. (1993). *Understanding comics: The invisible art* (1st ed.). Harper Perennial.

McMaster, K.L., van den Broek, P, Espin, C.A., White, M.J., Rapp, D.N., Kendeou, P.., Bohn-

Gettler, C.M., & Carlson, S. (2012). Making the right connections: Differential effects of reading intervention for subgroups of comprehenders. *Learning and Individual Differences, 22*, 100-111.

McNamara, D. S., & Magliano, J. P. (2009). Towards a comprehensive model of comprehension. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 51, pp. 297–384). Elsevier Science.

McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction, 14*(1), 1-43.

Meringoff, L.K. (1980). Influence of the medium on children's story apprehension. *Journal of Educational Psychology, 72*(2), 240-249.

Millis, K. & Larson, M. (2008). Applying the Construction-Integration Framework to aesthetic responses to representational artworks. *Discourse Processes, 45*, 263-287.

Narvaez, D., van den Broek, P., & Ruiz, A. B. (1999). The influence of reading purpose on inference generation and comprehension in reading. *Journal of Educational Psychology, 91*(3), 488–496. https://doi.org/10.1037/0022-0663.91.3.488

Ness-Maddox, H., Dahl, A.C., Tadeo, M., Troemel, V., Berry, J., Carlson, S.E., & Magliano, J.P. (July 2022). Emotional Inferences Due to Framing and Composition During Think-Aloud Task with Non-Linguistic Graphic Narratives. Standard presentation at the annual meeting for the Society for Text and Discourse, Atlanta, Georgia. [International]

Pellicer-Sánchez, A., Tragant, E., Conklin, K., Rodgers, M., Serrano, R., & Llanes, À. (2020).

Young learners' processing of multimodal input and its impact on reading comprehension: An eye-tracking study. *Studies in Second Language Acquisition, 42*(3), 577–598. https://doi.org/10.1017/S0272263120000091

Poulsen, D., Kintsch, E., Kintsch, W., and Premack, D. (1979). Children's comprehesnion and memory for stories. *Journal of Experimental Child Psychology*, *28*(3), 379-403.

Renz, K., Lorch, E. P., & Milich, R. (2003). On-Line Story Representation in Boys With Attention Deficit Hyperactivity Disorder. *Journal of Abnormal Child Psychology, 31*(1), 93–104.

Rumelhart, D. E. (1977). Understanding and summarizing brief stories. In D. LaBerge & J. Samuels (Eds.), *Basic processes in reading: Perception and comprehension*. Lawrence Erlbaum Associates.

Schmidt, C.R., Paris, S.G., and Stober, S. (1979). Inferential distance and children's memory for pictorial sequences. *Developmental Psychology*, *15*(4), 396-405.

Schwabe, A., Lind, F., Kosch, L., & Boomgaarden, H. G. (2022). No negative effects of reading on screen on comprehension of narrative texts compared to print: A meta-analysis. *Media Psychology*. https://doi.org/10.1080/15213269.2022.2070216

Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly, 24*, 402-433.

Stein, N. L., & Glenn, C. G. (1979). An analysis of story comprehension in elementary school children. In R.O. Freedle (Ed.), *Advances in discourse processes (Vol. 2): New directions in discourse processing.* Ablex.

Stracyznski, J.M., Romita Jr., J., and Hanna, S. (2002, February). "Meanwhile…" *Amazing Spider-Man, 2*(39). [Comic book]. Marvel Comics.

Tompkins, V., Guo, Y, and Justice, L.M. (2012). Inference generation, story comprehension, and language skills in the preschool years. *Read Writ, 26*, 403-429.

Trabasso, T. & Suh, S. (1993). Understanding text: Achieving explanatory coherence through on-line inferences and mental operations in working memory. *Discourse Processes, 16*(1-2), 3-34. DOI: 10.1080/01638539309544827

Trabasso, T., & Magliano, J. E. (1996). Conscious understanding during text comprehension. *Discourse Processes, 21*, 255–288.

Trabasso, T., van den Broek, P., & Suh, S. (1989). Logical necessity and transitivity of causal relations in stories. *Discourse Processes, 12*, 1–25.

van den Broek, P., Beker, K., & Oudega, M. (2015). Inference generation in text comprehension: Automatic and strategic processes in the construction of a mental representation. In E. J. O'Brien, A. E. Cook, & R. F. Lorch, Jr. (Eds.), *Inferences during reading* (pp. 94–121). Cambridge University Press.

van den Broek, P., Lorch, R., Linderholm, T., & Gustafson, M. (2001). The effects of readers' goals on inference generation and memory for texts. *Memory & Cognition, 29*, 1081–1087. doi: 10.3758/bf03206376

van Dijk, T., & Kintsch, W. (1983). *Strategies of discourse comprehension.*. Academic Press.

van Neste, J., Hayden, A., Lorch, E.P., & Milich, R. (2015). Inference generation and story comprehension among children with ADHD. *J Abnorm Child Psychol, 43*, 259-270.

Wannagat, W., Waizenegger, G., and Niedling, G. (2017). Multi-level mental representations of written, auditory, and audiovisual text in children and adults. *Cogn Process, 18*, 491-504.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001; 2014). *Woodcock Johnson III Tests of Achievement.* Riverside Publishing

Wright, H.H., Capilouto, G.J., Srinivasan, C., & Fergadiotis (2011). Story processing ability in cognitively healthy younger and older adults. *Journal of Speech, Language, and Hearing Research, 54,* 900-917.

Yeari, M., Elentok, S., & Schiff, R. (2017). Online and offline inferential and textual processing of poor comprehenders: Evidence from a probing method. *Journal of Experimental Child Psychology*, *155*, 12–31. https://doi.org/10.1016/j.jecp.2016.10.011

Zacks, J. M., Speer, N. K., & Reynolds, J. R. (2009). Segmentation in reading and film comprehension. *Journal of Experimental Psychology: General*, *138*(2), 307–327. https://doi.org/10.1037/a0015305.supp (Supplemental)

Zhang, M., Chen, Y., Lin, Y., Ding, H., & Zhang, Y. (2022). Multichannel perception of emotion in speech, voice, facial expression, and gesture in individuals with autism: A scoping review. *Journal of Speech, Language, and Hearing Research, 65*(4), 1435–1449. https://doi.org/10.1044/2022_JSLHR-21-00438

Zwaan, R., Langston, M., & Graesser, A. (1995). The Construction of Situation Models in Narrative Comprehension: An Event-Indexing Model. *Psychological Science, 6*(5), 292-297. https://doi.org/10.1111/j.1467-9280.1995.tb00513.x

APPENDICES

Appendix A

Practice Text Narrative

Jimmy wanted a motorcycle.

His mom suggested that Jimmy find work to afford the bike.

Jimmy took a part-time job at the grocery store.

At the end of the summer, Jimmy picked out a blue motorcycle.

Practice Graphic Narrative

Narrative A Text Version

FK 5.2

Mary Jane lay on one side of the bed alone, the other side of the bedding barely pulled back.

She faced the alarm clock which read 5:29 A.M.

Mary Jane frowned.

She got up and slipped off the shoulder strap of her nightgown as she rose.

Mary Jane walked into the bathroom as the TV played a commercial.

A newscaster came on the TV and reported from inside a studio.

The camera switched to a reporter outside in the field.

Footage played of Spider-Man swinging by and saving people from a man firing a gun.

Mary Jane emerged from the bathroom and stood in front of the TV as she watched the report on Spider-Man.

On the news, Spider-Man carried people out of harm's way, and Mary Jane continued to watch.

The news switched to a frankfurter commercial, and Mary Jane returned to the bathroom.

Mary Jane modeled a bathing suit on the beach in front of a photographer and his support crew.

She smiled and posed while flashbulbs went off.

Mary Jane smiled as she emerged from a limo.

She moved down the red carpet, smiling and waving.

People behind and beside her laughed uproariously at a movie screen, but Mary Jane looked down, not laughing or smiling.

She emerged from the theater, smiling and blowing a kiss.

Her limo pulled away from the curb.

Back in her bedroom, she was getting ready for bed while the TV played a hospital drama.

While Mary Jane was still in the bathroom, the TV played a mouthwash commercial.

A Spider-Man news story came on the TV.

Mary Jane came out of the bathroom and watched the story.

She leaned against the doorframe while watching the TV.

She got into bed on just the one side.

She lay on one side of the bed, looking at the alarm clock.

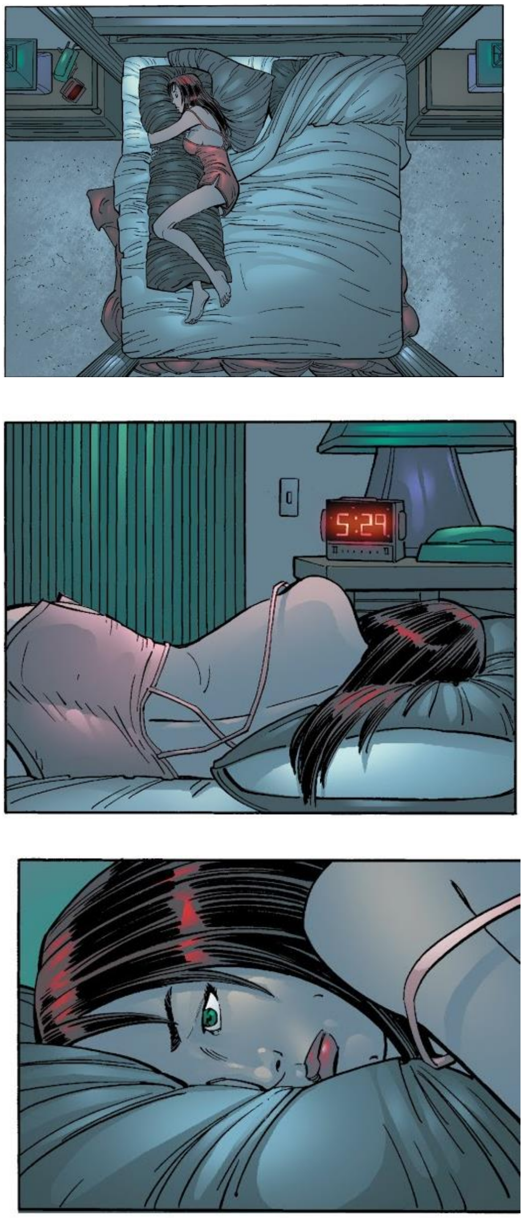
The alarm clock read 11:30 P.M.

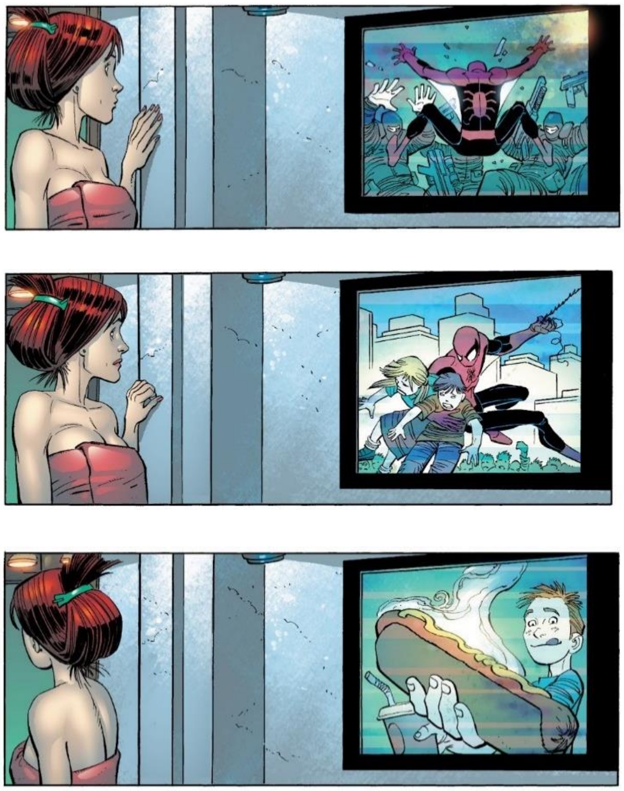
Mary Jane lay awake.

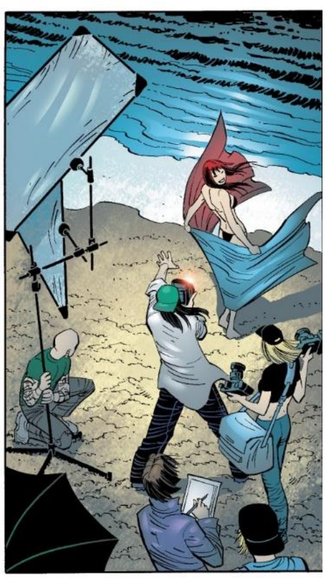
The alarm clock read 5:29 A.M.

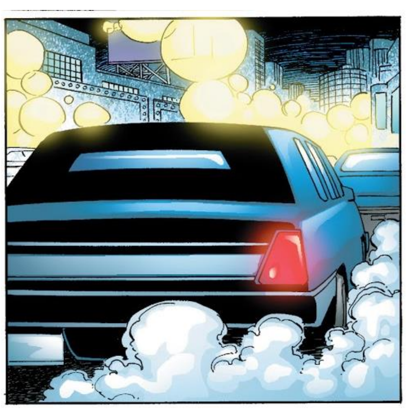
Mary Jane lay awake and looked at the unoccupied pillow on the other side of the bed.

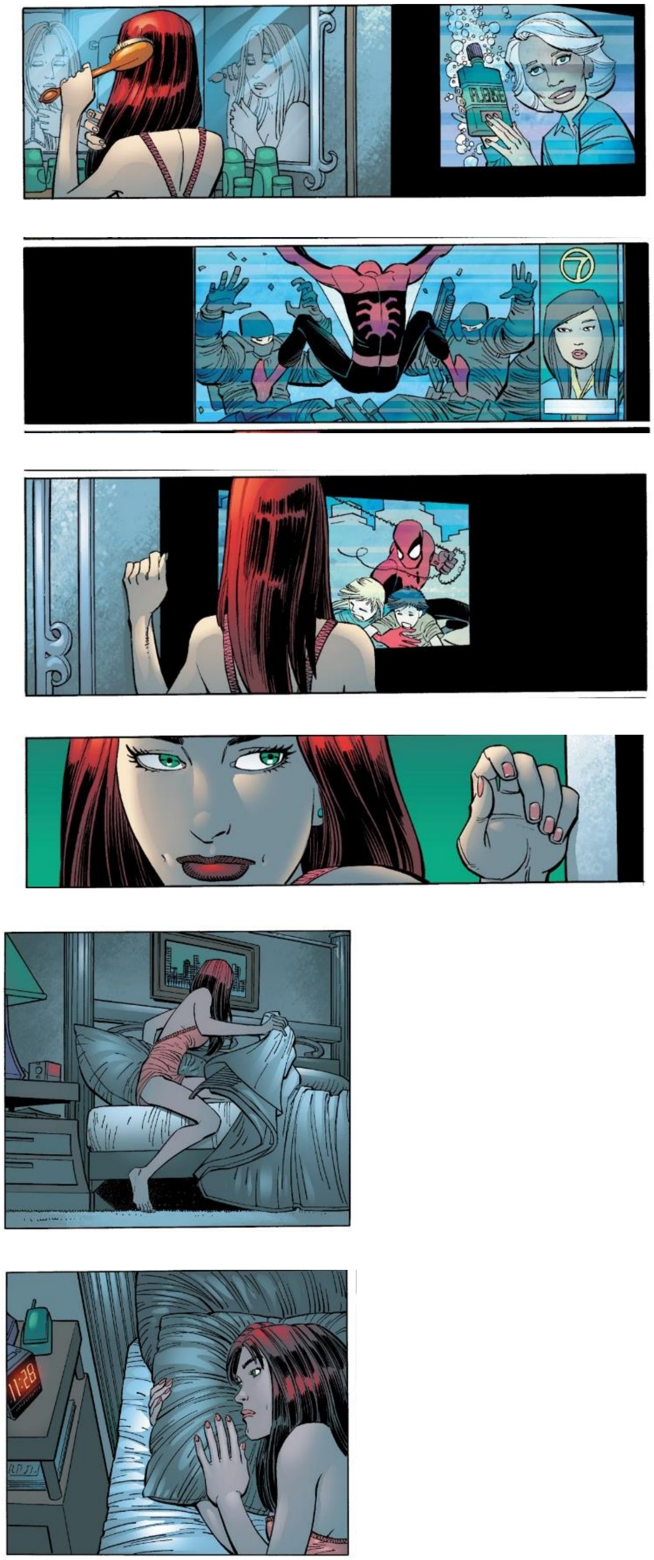
Narrative A Graphic Version










11:28


11:30
5:29

Narrative B Text Version

FK 4.0

Lloyd held his head in his hands.

He sat at his desk.

He picked up a framed photo on his desk.

The photo showed a woman and two young children, a boy and a girl.

He placed it in his box right on top.

Lloyd exited his small, shabby office with his box of stuff.

An older woman wiped away tears and talked to Lloyd.

Lloyd looked down the hallway to a number of employees gathered around a Christmas tree, drinking, partying, and celebrating.

Lloyd frowned as the image of partiers reflected in his glasses.

Lloyd carried his box of possessions out of the office building.

Lloyd walked down the street while festive people surged by, many carrying packages.

Lloyd stopped in front of a jewelry store with a display of jewelry in the window.

He gazed at an ornate necklace.

Lloyd turned to look over his shoulder.

A donation worker rang a bell while standing next to a collection pot.

Lloyd took out his wallet from his coat pocket.

He opened it to reveal a single five-dollar bill.

Lloyd frowned down at his wallet.

He dropped the single bill into the collection pot as he strolled by.

Lloyd pulled into the driveway of a small, simple house.

Lloyd walked up to the house with his box of stuff.

He stood on the front step, looking in the window where children were playing on the living room floor.

He held his head in his hand while snow fell around him.

He bent down and hid the box behind a bush next to the steps.

Lloyd walked through the door, smiling, greeted by the boy and girl.

He walked into the kitchen with the two children in his arms where a woman sat the kitchen table pouring over papers.

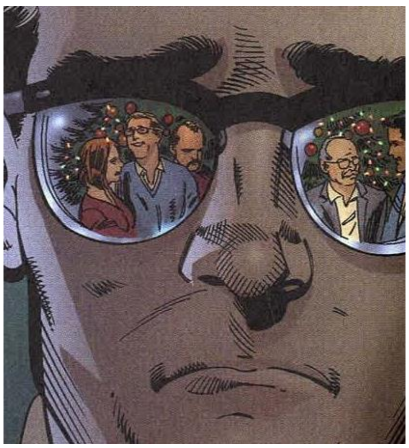
She held up bills in each hand.

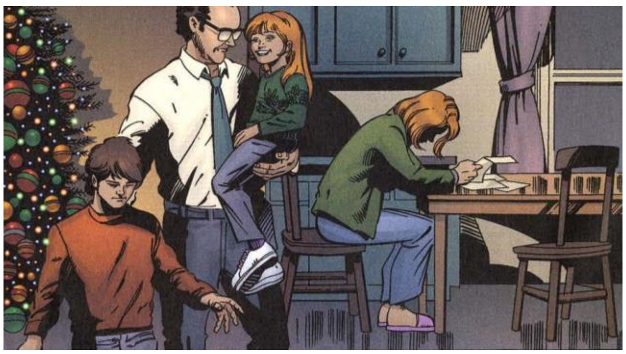
Lloyd hustled the children out of the kitchen.

He placed a hand on her shoulder while she lowered her head in her hands and cried.

She pointed, and she and Lloyd looked to the tree with nothing under it.

Narrative B Graphic Version






OVERDUE

Narrative C Text Version

FK 4.4

Bruce Banner stepped inside a roadside diner.

Banner seated himself in a booth by the window.

As he sat down, he took off his coat.

A mother and her young daughter sat in the booth opposite Banner, the girl's pupils big as she stared.

He smiled and wiggled his fingers.

Banner set up his laptop.

He looked up and sipped his coffee.

A black sedan pulled into the lot.

Three men in black suits and black sunglasses emerged from the car.

The men in black headed into the diner through the screen door, one of them accidentally stepping on the tail of a lop-eared mutt as they passed.

Banner watched them enter the diner.

The men in black seated themselves at the counter across from Banner, and Banner raised his hand to hide his face.

The waitress poured the men in black coffees.

Banner threw some cash onto the counter and left his booth.

He passed the men in black and pushed open the door to the restroom.

A man in black looked up, and Banner's face reflected in the lenses of his sunglasses.

Banner rested against the door in the little restroom, with one small, high window above the toilet.

In the diner, one of the men in black pointed to the bathroom, gesturing to his colleagues.

One of the men in black got up and walked to the restroom, hand in his blazer.

The man in black entered the restroom.

Gun drawn, he spotted the small high window, broken open!

Looking out the window, he motioned to his companions toward outside the diner.

The rest of the men in black left the counter.

They ran to out back of the diner, having drawn their own weapons.

The leader signaled to the others to scour the area.

Inside the restroom was a small storeroom beside the toilet.

Banner stepped out of the small storeroom.

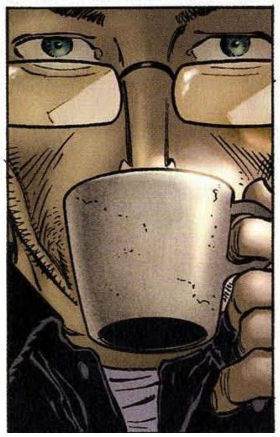
He peered through the restroom door.

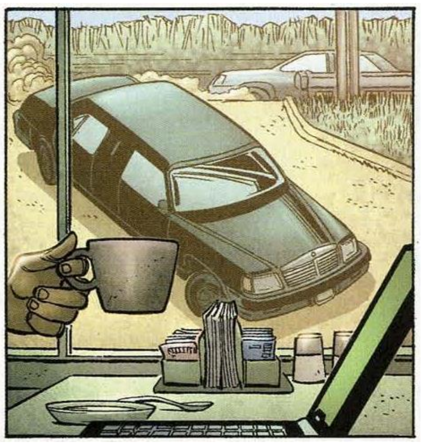
Banner moved through the diner.

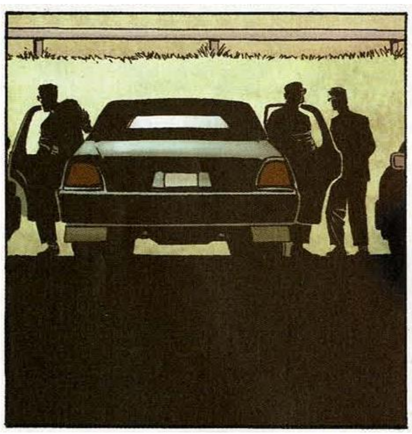
Banner stepped through the screen door, greeted by no one but the lop-eared mutt.

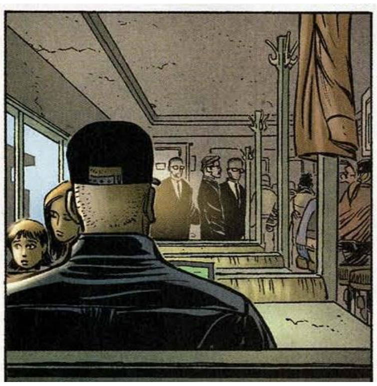
He walked into the cornfield behind the diner.

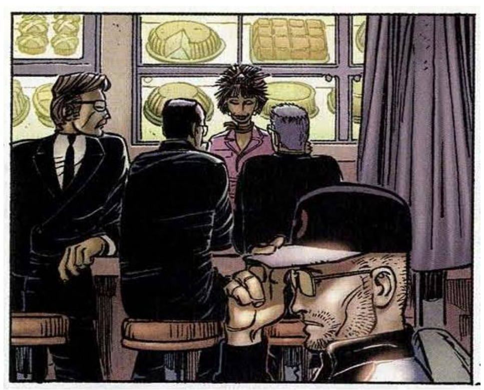
Narrative C Graphic Version












RESTROOM


RESTROOM


RESTROOM


CLOSED

Narrative D Text Version

FK 4.4

Young Thor and Loki played in the castle garden with wooden swords, Thor's sword broken from Loki's swing.

Thor looked at his sword, frowning, and Loki gestured for Thor to follow him.

Thor followed Loki through a small back door of the castle.

Within the castle, Odin and a group of warriors celebrated a tremendous victory by raising tankards of ale, gnawing on turkey legs, and throwing darts; a bow and quiver leaned against a chair.

Thor and Loki looked down on the celebration from the rafters.

Loki pointed at Thor, but Thor raised his hands and shook his head.

Grinning, Loki presented a small gold coin taken from under his belt.

Loki flipped the coin in the air.

Loki presented the coin heads side up.

Thor frowned.

Loki remained in the rafters, while Thor scrambled off.

Thor crawled through the warriors' legs toward the bow and quiver.

Thor raced out the doorway of the chamber, all the warriors of Asgard not noticing that he had the bow and quiver in his arms.

At the edge of the garden, Thor and Loki shot arrows toward one tree that bordered the woods. Loki had the bow.

Loki handed the bow over to Thor.

As Thor drew back the bowstring, Loki laughed at his struggle and teased Thor.

Sticking out his tongue and closing one eye, Thor pulled back the string a bit more.

As he laughed, Loki slapped Thor's shoulder, causing Thor to let go of the string.

The arrow flew off the bow.

The arrow flew past the tree, a beautiful deer in its path.

Loki and Thor stared in the direction of the misfired arrow.

Loki turned and fled, while Thor walked toward the woods.

Thor knelt in front of the deer with an arrow stuck in its hide.

Tears ran down Thor’s face.

Back in the castle, Odin was telling the warriors a story, but their attention was diverted elsewhere.

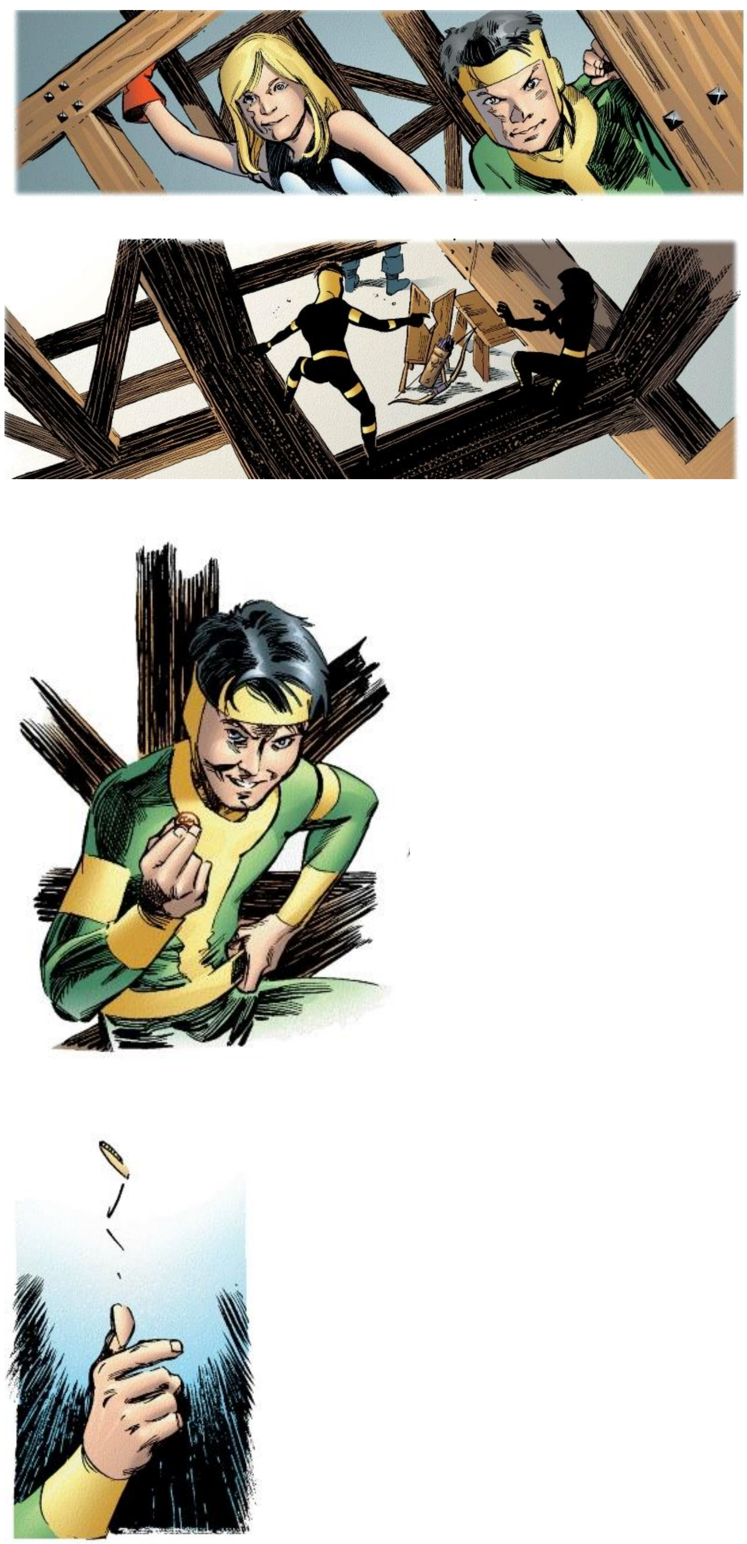
Odin turned to see a stoic Thor walk into the room with the bow and quiver.

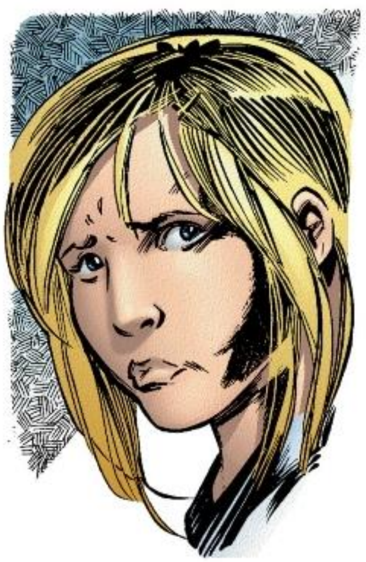
Thor broke down and sobbed.

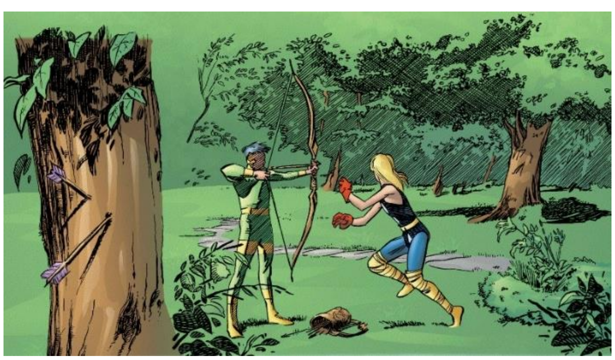
Odin placed his hands on Thor’s shoulders.

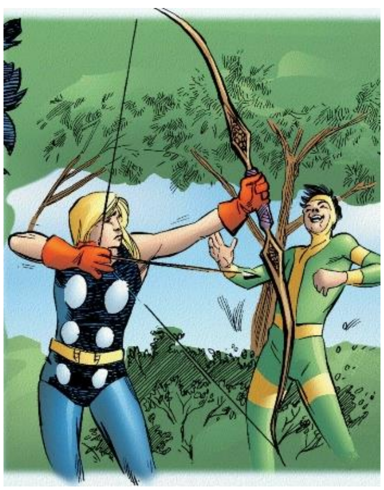
Odin knelt and hugged Thor.

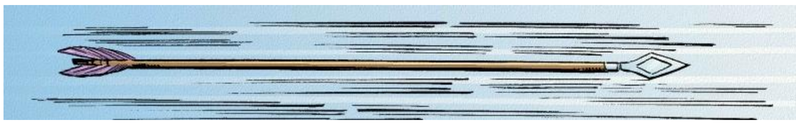
Narrative D Graphic Version

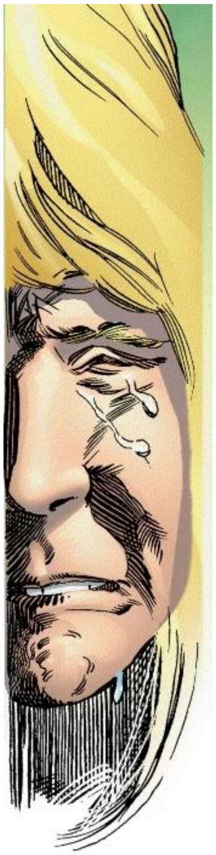












Appendix B

Visual Language Fluency Index (Cohn, 2014)

1. Using the following scale, on average, how often per week do/did you…? (place a whole number in the square)

Never ---------------------------------Sometimes---------------------------------Always
1 2 3 4 5 6 7

	Currently (mark 1 – 7)	While growing up (mark 1 – 7)	Which is your favorite? (mark with "X")
…read text-only books for enjoyment			DO NOT MARK HERE
…watch movies			DO NOT MARK HERE
…watch cartoons/anime			DO NOT MARK HERE
…read comic books			
…read comic strips			
…read graphic novels			
…read Japanese comics (manga)			
…draw comics			

2. How would you rate your expertise with reading comics (of any sort)? (Mark "X" once in each row)

	Above average (5)	Slightly above average (4)	Average (3)	Slightly below average (2)	Below average (1)
Currently					
While growing up					

3. How would you rate your drawing ability? (Mark "X" once in each row)

	Above average (5)	Slightly above average (4)	Average (3)	Slightly below average (2)	Below average (1)
Currently					
While growing up					

4. How old were you when you began reading comics? ___________ Drawing comics? ____________

Author Recognition Test (Acheson et al., 2008)

Subject Number:______

Score: C_____I_____C-I_____

Below is a list of names. Some of them are authors of books, and some of them are not. Please put a check mark next to the ones that you know for sure are authors. There is a penalty for guessing, so you should check only those names about which you are absolutely certain. Thank you.

___Patrick Banville
___Kristen Steinke
___Ernest Hemingway
___Clive Cussler
___Hiroyuki Oshita
___Kurt Vonnegut
___Anne McCaffrey
___Elinor Harring
___Sue Grafton
___Lisa Woodward
___David Harper Townsend
___Anna Tsing
___T.C. Boyle
___Jonathan Kellerman
___Cameron McGrath
___F. Scott Fitzgerald
___A.C. Kelly
___Peter Flaegerty
___Kazuo Ishiguro
___Jane Smiley
___James Patterson
___Martha Farah
___Craig DeLord
___Nora Ephron
___Ann Beattie
___Stewart Simon
___Danielle Steel
___Dick Francis
___Ted Mantel
___I.K. Nachbar
___Judith Krantz
___Thomas Pynchon
___Wayne Fillback

___Harry Coltheart
___Gary Curwen
___Herman Wouk
___Geoffrey Pritchett
___Ray Bradbury
___Jay Peter Holmes
___Christina Johnson
___Jean M. Auel
___Judith Stanley
___Gloria McCumber
___James Joyce
___Robert Ludlum
___Larry Applegate
___Keith Cartwright
___Jackie Collins
___Umberto Eco
___David Ashley
___Jack London
___Seth Bakis
___Padraig O'seaghdha
___E.B. White
___Giles Mallon
___Raymond Chandler
___Isabel Allende
___Amy Graham
___Marion Coles Snow
___George Orwell
___Maya Angelou
___Bernard Malamud
___John Grisham
___Erich Fagles
___Walter Dorris
___Gabriel Garcia Marquez

___Virginia Woolf
___John Landau
___Toni Morrison
___Harriet Troudeau
___Roswell Strong
___J.R.R. Tolkien
___Margaret Atwood
___Seamus Huneven
___Harper Lee
___Chris Schwartz
___Walter LeMour
___Alice Walker
___Elizabeth Engle
___T.S. Elliot
___Marvin Benoit
___Joyce Carol Oates
___Jessica Ann Lewis
___Nelson Demille
___Arturo Garcia Perez
___S.L. Holloway
___John Irving
___Stephen Houston
___Marcus Lecherou
___Valerie Cooper
___Tom Clancy
___Vladimir Nabokov
___Pamela Lovejoy
___Vikram Roy
___Saul Bellow
___Stephen King
___Elizabeth May Kenyon
___Frederick Mundow

___Tony Hillerman
___Amy R. Baskin
___James Clavell
___Salmon Rushdie
___Maryann Phillips
___Scott Alexander
___Ayn Rand
___Alex D. Miles
___Margaret Mitchell
___Leslie Kraus
___Ralph Ellison
___Sidney Sheldon
___ Brian Herbert
___Sue Hammond
___Jared Gibbons
___Michael Ondaatje
___Thomas Wolfe
___Jeremy Weissman
___Willa Cather
___J.D. Salinger
___ Antonia Cialdini
___ Lisa Hong Chan
___Samuel Beckett
___Beatrice Dobkin
___Wally Lamb
___Katherine Kreutz
___James Michener
___William Faulkner
___Isaac Asimov
___Lindsay Carter
___Paul Theroux
___Francine Preston

Demographic Survey

1. What is your age? ______________

2. What is your gender identity?
    a. Man
    b. Woman
    c. Trans man
    d. Trans woman
    e. Non-binary / gender fluid
    f. Agender

3. Which of these options match your racial identity? Check all that apply.
    a. American Indian
    b. Asian
    c. Black
    d. Pacific Islander
    e. White
    f. Other
    g. Prefer not to answer

4. What is your ethnicity?
    a. Hispanic
    b. Non-Hispanic

5. What is your class standing?
    a. Freshman
    b. Sophomore
    c. Junior
    d. Senior / Graduate

6. Are you an international student?
    a. Yes
    b. No

7. How would you rate your CURRENT expertise with reading text narratives (of any genre) outside of school for your own enjoyment? Choose one.
    a. Above average
    b. Slightly above average
    c. Average
    d. Slightly below average
    e. Below average

8. How would you rate your expertise WHILE GROWING UP with reading text narratives (of any genre) outside of school for your own enjoyment? Choose one.
    a. Above average
    b. Slightly above average
    c. Average
    d. Slightly below average
    e. Below average

9. How would you rate your CURRENT text narrative writing ability? Choose one.
    a. Above average
    b. Slightly above average
    c. Average
    d. Slightly below average
    e. Below average

10. How would you rate your text narrative writing ability WHILE GROWING UP? Choose one.
    a. Above average
    b. Slightly above average
    c. Average
    d. Slightly below average
    e. Below average

11. How old were you when you began reading text narratives for enjoyment? (Type N/A if not applicable) ___________

12. How old were you when you began writing text narratives? (Type N/A if not applicable) ___________