Creating Comics from Automatic Summarization of Sports Stories

by

TAO LI

A thesis submitted to
The University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

Department of Electronic and Electrical Engineering
The University of Birmingham
October 2023

Abstract

Automatic text summarisation offers a means of managing information overload by presenting the critical points in documents. This thesis proposed to improve this process by visualising such summaries. Sports match reports from websites are summarised in the form of comics. Instead of using machine learning to support summarisation, we used the story grammar to support automatic summarisation. We will define a sports match's 'story grammar' (in terms of events, such as scoring a goal, penalising a foul, etc., and Named Entities, such as teams, players, referees etc.). It frames the extractive-summarisation process. Then we visualised the summary into a comic. We conducted two user trials to evaluate the summary. In the first user trial, participants read reports of soccer games in 3 different conditions. Then they answered comprehension questions. The results showed that the participants read the comics more quickly than the text, resulting in superior comprehension (literal, inferential, and evaluative), and the participants preferred to choose comics as reading material. In the second study, we designed an eye-tracking experiment to explore reading strategies for different media. While reading time is faster for comics, the average fixation time per word is the same across all media. It suggests the texts were skim-read, which might explain the lower comprehension of the text. Surprisingly, the comic is highly supportive of evaluative and inferential comprehension. Overall, the comic produced significantly better performance on comprehension tests in both studies. As we hypothesised, comics could significantly reduce time consumption and improve comprehension. Therefore, we believe that the comic helps users access important information.

Acknowledgement

I would like to express my profound gratitude to Professor Chris Baber, an unwavering supervisor dedicated to education. His meticulous guidance and invaluable assistance in my research were exceptional. His generous investment of time, energy, and the unwavering respect he extended to me have profoundly touched my heart. In addition, my gratitude extends to Sagir, Hebah, and all others who lent their expertise to my dissertation. Your invaluable contributions were instrumental in the realization of this project. I am deeply indebted to my beloved parents and my wife, Hongyu, who steadfastly stood by my side throughout my doctoral journey. Their assistance in navigating various life and emotional challenges was indispensable. The depth of their love and support provided a solid foundation during times of distress. Furthermore, I wish to extend my gratitude to the more than 80 friends who actively participated in the experimental study. Their enthusiastic support played a pivotal role in ensuring the smooth progression of my research. Finally, my heartfelt thanks go to the University of Birmingham for providing an excellent research environment. This institution enabled me to carry out my work seamlessly, and I take immense pride in having been part of an institution that upholds high academic and research standards.

Contents

1 Introduction

1.1 Aims and objectives

This PhD thesis aims to summarise and illustrate soccer game text reports using comics. The assumption is that comics support rapid information acquisition while maintaining clarity and critical information. In other words, the information contained in comics are easier to understand than text reports. Furthermore, most panels of comics can be displayed in a single window to reduce scrolling or other interaction. The research proposes a prototype system which produces summaries of soccer game reports and renders these in the form of simple comic strips. The next step is user trials, which measure whether there is a noticeable difference in comprehension when users read original reports, text summaries, and comics.

1.2 Research questions

- What is the most appropriate way to summarise text to render information in the form of comics?
- Is there any difference between the original text, abstract text and comic material for readers?
- Which of the three materials, the original text, the summary text, and the comic material, can best help readers understand the events described?

1.3 Background

We receive a bombardment of information through numerous digital platforms in today's age. This information overload and data bloat drain human resources and reduce the efficiency of information extraction further (Gambhir and Gupta, 2017). In the multimedia age, text, pictures, and videos have evolved into essential carriers of information. However, the increasing quantity of text information might make it challenging for people to extract relevant information efficiently. This challenge has spurred growth in the interest in techniques for summarising and visualising information from text. The design of automatic text summarization systems requires the use of cumbersome and complex techniques to manage and simplify text complexity (Kucher and Kerren, 2015). The most fundamental purpose of automatic text summarization systems is to reduce the text volume while retaining meaningful content (Gambhir and Gupta, 2017). Alongside developments in text summarisation, a trend to use visualisation to present the content of text has emerged. For example, you can manually summarise texts in the form of word clouds or mind maps, tables, lists, or arrow diagrams (Dymock, 2005). Furthermore, we can summarise information in the form of dashboards (e.g., Tableau or d3) or as layers of temporal, spatial, and social network data (Eccles et al., 2007). This thesis has a particular interest in using comics as a popular form of visualisation.

One of many benefits of comics involves using pictures to convey information efficiently. Pictures have proven vital in early childhood education because they vividly express relevant information (Lin et al., 2015). Interesting pictures can significantly stimulate children's desire to learn and explore and can assist children in establishing independent learning (Lin et al., 2015). One might suggest that comics can effectively encourage readers to understand stories, as using minimal text and supporting images can express the underlying message (Lin et al., 2015). Research on the effectiveness of comics for learning is limited. In one study (Lin et al., 2015), comics appeared to hinder the performance of more able students, showed little effect on less able students, and provided some benefits for students in the class. One reason for these differences could relate to the "learning" measures for these students. We can measure learning by comprehension tests and define it in literal, reasoning, and evaluation terms (Basaraba et al.,2013). Literal comprehension involves extracting and recalling superficial details, while inferential comprehension involves relating content to prior knowledge. Evaluation comprehension requires the reader to reason "beyond" the text. We might assume that comics are particularly effective in subjects that require people to learn and remember "facts," but less effective in situations where they may need to reason beyond those facts. On the other hand, comics have demonstrated a means of exploring "difficult" social and moral issues (Versaci,2001). For example, creators have used comics to explore the experience of HIV/AIDS (Czerwiec,2018). Regarding this example, comics' role was not so much to present facts but to convey experiences in a way that readers can empathize with and reflect on. From this perspective, the role of comics might involve less presenting of facts and more encouraging readers or learners to engage with the material. In this thesis, the focus will be on the reader's comprehension in terms of their ability to recall information relevant to a specific topic. The thesis will use reports of soccer games as the use-case, and the question is whether presenting information related to soccer in the form of a comic that has been automatically generated from a summary of the report produces comprehension comparable to reading the full report. This process requires the ability to convert text directly into pictures. We will describe the system developed to accomplish this in chapter 2. But before doing so, we will consider the ways in which people read text and comics.

1.3.1 Theories of reading text

Reading text-based communications has been a human activity for thousands of years. Through written words, writers can communicate with others at a distance. Readers can perceive arbitrarily determined information presented in specific contexts and then translate them into meaning (Hudson,1998). Imagine that you're about to read a conference paper. This document contains several elements, e.g., titles, abstracts, subtitles, paragraphs, etc. You can logically connect these elements into a hierarchy (figure 1).

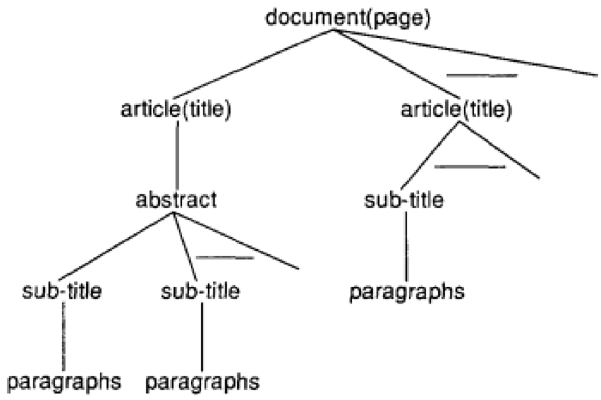


Figure 1. The construction of one document(Tsujimoto and Asada,1992)

Reading a document involves extracting information from the lowest levels of the hierarchy (e.g., at the level of the words that make up the paragraphs) to extract the most important information. Although we can glean a certain amount of important information from abstracts and titles that have already been processed, this summarized information only contains concepts to a certain extent. Therefore, this might prevent a reader from obtaining more detailed information. We can combine the important information into a gist (Schank, 1992) of the document. The gist is a running summary of the document that you build up during the reading process. Our prior knowledge and the capacity of your working memory will limit this process. For example, if we know a lot about the material in the document, the gist will involve activating knowledge we already have. But if the material is new, creating the gist will involve a great deal of effort in relating the new information to your prior knowledge. Traditionally, we relate such 'gist' to the mental model (Johnson-Laird et al.,1992) or schema (Bartlett, 1932) that we construct as we read the document.

At the word level, reading involves moving your eyes to extract information from words on the page. Eye movements involve fixations (where the eye rests on individual letters) and saccades (where the eye quickly moves between letters). As figure 2 suggests, we don't need to make fixations on every letter. We can skip some letters or even some words (usually because you have assumed what the word might be), or some movements might be backwards (regression) in the text, e.g., from 7 to 6 in figure 2. We define an area of interest when several fixations occur in close proximity.

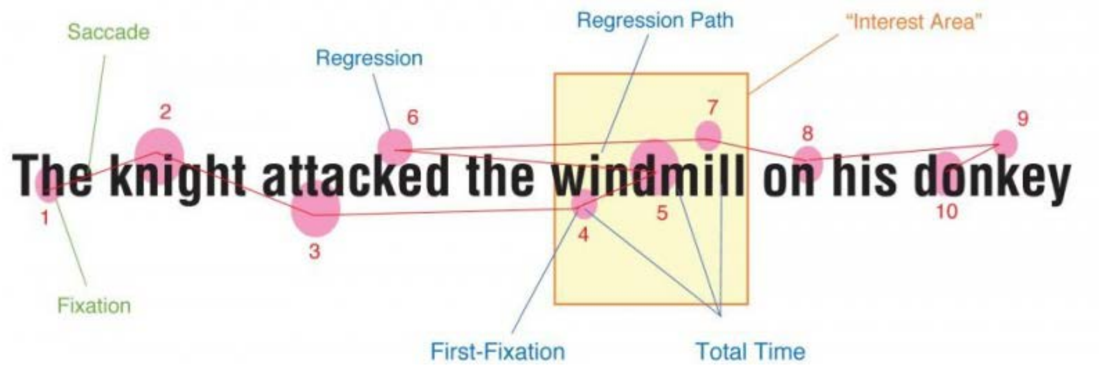


Figure 2. Reading and eye tracking [American Academy of Ophthalmology, 2020. 'Saccades and fixations'. Available at: https://eyewiki.aao.org/File:AA0_61866.jpg (Accessed: May 27, 2020)].

Contemporary theories of reading tend to combine top-down processes, where the reader anticipates the likely next words in a sentence, and bottom-up processes, where the reader uses the information from fixations to define letters and words. The reader guides their gaze along the line of text in a top-down, word-based manner. Younger readers generate more regular left-to-right directional scan paths, while older readers show more regression (von der Malsburg et al., 2015; Albrengues et al., 2019). Complicating matters further, the bottom-up processes seem to also involve top-down processes. For instance, the fixations might involve not just individual letters but also the shapes (or envelope) of words, allowing you to read a word even if the letters are scrambled (the word "rghit" could read as 'right'). The space between words, which is the most prominent visual feature in most modern texts, also influences saccades. Experiments on reading unspaced text show a modest increase in fixation duration and a decrease in overall reading speed (Epelboim, Booth and Steinman,1994). The research found that the length of progressive (right) saccades was the only global eye movement aspect that changed significantly when the text had spaces (Epelboim, Booth and Steinman,1994).

As readers become more skilled in reading and use it in an expanding range of contexts, the processes of interpreting letters and words become increasingly automatic. Bottom-up and top-down approaches result from views on the nature of reading that depend on scale. As previously stated, most researchers favor an interactive approach. The information processing and integrative analysis paradigm developed by Gibson and Levin (1975) forms the foundation for two additional reading strategies. Both strategies expose to the general public the relationship between readers and texts. The bottom-up approach suggests that readers create meaning from letters, words, phrases, clauses, and sentences by processing the text sequentially into phonemic units expressing lexical meaning and then linearly creating meaning. This approach focuses on cognitive information processing difficulties. This method presumes that the reading work divides into several stages for understanding related tasks that progress in a predetermined order from sensory input to comprehension. Additionally, visual language processing is the most critical factor in reading comprehension, as opposed to auditory comprehension (Hoover and Tunmer, 1993; Gough, 1972). Typically, researchers expect very passive, fast, and efficient information gathering. Likewise, the impact of previously comprehended and stored knowledge on current information processing is negligible (Hudson,1998).

Carroll's (1964) definition of reading captures the assumption in traditional bottom-up approaches:

“Reading is the reconstruction (overt or covert) of plausible oral information from a printed text, and the meaningful interpretation of the reconstructed information. Response activities would be similar to those that would have had on spoken messages" (Carroll, 1964).

The concepts of phoneme-glyph correspondence and reconstructing the information processing view of existing messages are crucial. The most typical example of a bottom-up strategy is "one-second reading" by Gough (1972). Gough sees the reader as a visual system that scans a series of letters, based on a fairly specific model. Gough argues, "*there is no reason to reject the assumption that we do read word for word. In fact, the weight of the evidence persuades me that we do read sequentially from left to right*" (Gough,1972). Bottom-up reading methods focus on the problem of fast text processing and word recognition. They heavily emphasize the reader's ability to recognize words by immediately mapping the input to some independent representation in the mental lexicon. Typically, people regard this mapping as context-independent. In many cases, they view context-dependence as a reading tactic used by weak readers (Perfetti, 1995; Thompson et al., 1993). Poor readers also struggle to connect material to its immediate meaning, so they must rely on context as a compensatory method.

The top-down strategy assumes that the reader reads the text at a higher conceptualization level than the next level (such as paragraphs, sentences, or words within the text), and then delves into the text itself. These methods assume slower information processing circuits than bottom-up methods due to psychological limits on memory capacity and information storage speed. As a result, readers form shifting assumptions about incoming data. This type of reader uses schemas of form and content to create customized and context-sensitive meanings for texts (Hudson,1998). Strong forms of these theories assume that readers are not limited by the text but use text samples to confirm expectations about the text's message (Smith,2012). Reading is a psycholinguistic process that involves interaction between thinking and language (Goodman,1976). Due to their syntactic and semantic expertise, readers typically rely less on printed and phonetic texts. Goodman (1976) prescribes four reading processes: prediction, sampling, confirmation, and correction. Goodman's model allows the reader to use grapheme-to-phoneme correspondences, but it prioritizes the cognitive efficiency of relying on existing syntactic and semantic information. Readers make informed guesses about the text's meaning and then sample the material to confirm or refute their assumptions.

Reading is an active activity that educates the reader not only in the language, but also in how to interact with the language's internal concepts, the background of past experience, and the context of universal notions. An interaction or trade-off between visual and non-visual processes contributes to effective reading. These strategies provide the fewest cues necessary for a good guess. Smith (2012) acknowledged the use of prediction and context as mediating "memory bottlenecks" during reading in discussing the importance of short-term and long-term memory. Smith proposes that reading is intentional and selective in that readers focus only on what is essential for their current purpose. This viewpoint suggests that reading instruction should happen in contexts where the text can be understood, rather than focusing on isolated phoneme-grapheme correspondence activities. It also suggests a greater emphasis on extensive reading rather than constant reading by interrupting the reader with probing questions. According to Smith (2012), reading is intentional and selective. Therefore, they present the START Method (Students and Teachers Actively Read Text) as a means to improve reading comprehension instruction and student success. This innovative educational framework enhances reading comprehension performance and instruction by actively involving students in strategic reading during independent reading and modeling and enhancing the eight comprehension strategies during teachers' reading. By using this framework, teachers became more proficient in instructing students in reading comprehension, assessing their reading comprehension skills, and employing metacognitive comprehension strategies (Scharlach,2008). Teachers who use the START framework maximize the educational value of what currently happens during school hours without subtracting from essential instructional time or adding extra preparation time. The START Framework provides teachers with an effective educational framework for enhancing reading comprehension instruction, with clear goals for student performance and a thorough focus on science-based reading research (Scharlach,2008).

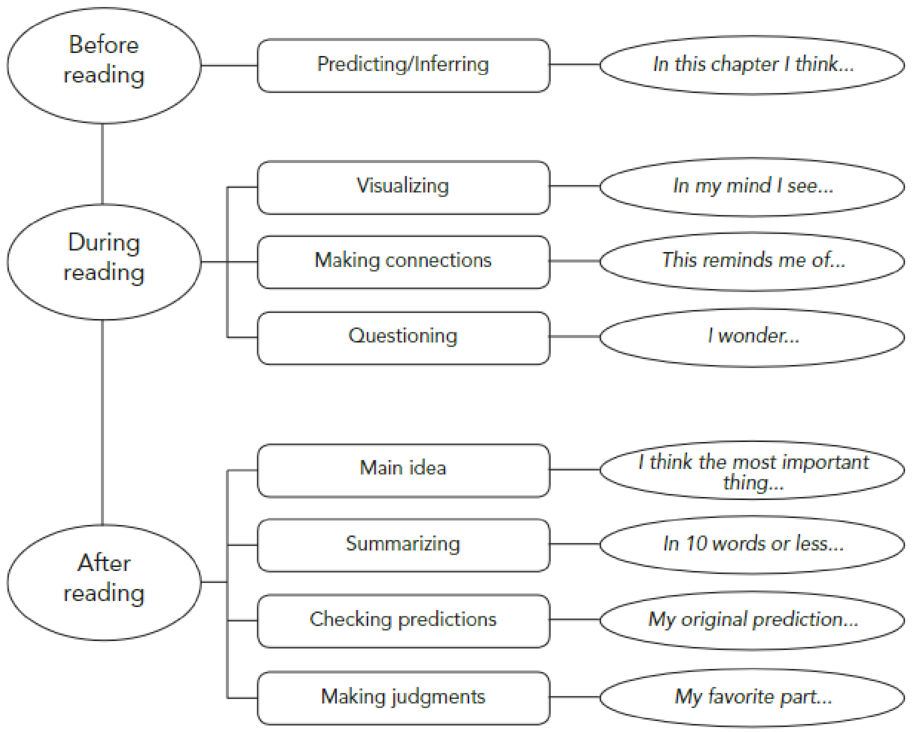


Figure 3. START Reading Strategies Diagram(Scharlach,2008)

From an interactive reading perspective, both the top-down and the bottom-up approaches appear simplistic. This perspective asserts that reading is fundamentally bidirectional and involves the use of prior knowledge, higher-order psychological processes, and text processing. This reading interaction theory prioritizes the steps of selecting a sample of text, speculating on its intended meaning, and then assessing it through further reading. One of the notable features of the interactive model is its ability to synthesize patterns while applying information from various sources of knowledge. These include feature extraction, orthographic knowledge, lexical knowledge, syntactic knowledge, and background (or semantic) knowledge. This approach indicates that the reader will rely on additional information sources to deal with more difficult material, regardless of the reader's reading proficiency (Hudson, 1998). While accurately determining the number of sources one uses can be challenging, other methods allow us to define how difficult a given text might be. One of the primary ways to do this is to determine the text's readability.

1.3.2 Defining Readability

Typically, a readability rating helps determine how easy or complex a paragraph or piece of text is to read. It is useful to know what educational level a reader requires to read a piece of literature without difficulty. According to a scale from 0 to 100, the Flesch Reading Difficulty Score indicates how understandable a section is (Flesch,1948). It shows how hard it is to comprehend the material. If the score is higher, the information will be easier to read and understand, and vice versa. For example, a material that received a score of 100 indicates that the material was the easiest to read, while a material that received a score of 0 indicates that the material was hardly understood. If the Flesch Reading Ease score is 70 or higher, it is considered to have an excellent Flesch reading score and to be relatively easy to read. The following equation determines the Flesch reading level:

206.835 - 1.015 × (total words ÷ total sentences) - 84.6 × (total syllables ÷ total words)[1].

Table 1. the level of Flesch Reading Ease score[1]

Score	Grade	Summary
90 - 100	5th grade	Very easy to read
80 - 90	6th grade	Easy to read
70 - 80	7th grade	Fairly easy to read
60 - 70	8th & 9th grade	Plain English
50 - 60	10th to 12th grade	Fairly difficult to read.
30 - 50	College	Difficult to read.
10 - 30	College graduate	Very difficult to read
0 - 10	Professional	Extremely difficult to read

[1] https://charactercalculator.com/flesch-reading-ease/

1.3.3 Reading Comics

Numerous documents throughout human civilization's history show that people have long used imagery to record and display knowledge. Campaigns now heavily rely on graphical abstractions, images, posters, movies, and other graphic elements to increase public engagement and facilitate effective communication. The use of graphic materials draws more attention and provides more room for detailed and comprehensible explanations suitable for specific audiences (Wang, 2020). Numerous studies have examined how the human brain interprets different types of images and written descriptions (Hegarty and Just, 1993; Holsanova et al., 2009; Rayner et al., 2001). However, these media typically present a single image. Visual stories, such as comics, depend on an ordered sequence of images. This use of sequentially arranged images resembles ‘reading' (Foulsham et al., 2016), where the reader's attention follows a path, much like the left-to-right reading of Roman script or the right-to-left reading of Arabic script. People often "read" when interacting with comics. In many cases, individuals must understand non-textual visual narratives. The "reader" must make sense of each section by reading the relevant panels in the order of their features. Finally, the "reader" must understand and consider the behaviors of these visual elements merged into a single event.

Typically, the process of understanding continuous visuals parallels that of understanding texts (Cohn, 2013c; Foulsham et al., 2016; McCloud, 1993). In comic book design, the creators must arrange the graphics in a reasonable and not too disorganized manner. The ultimate goal of comic design is to provide a method that minimizes the need for semantic thought and reasoning, allowing readers to quickly understand what the comic is conveying (Cohn, 2015).

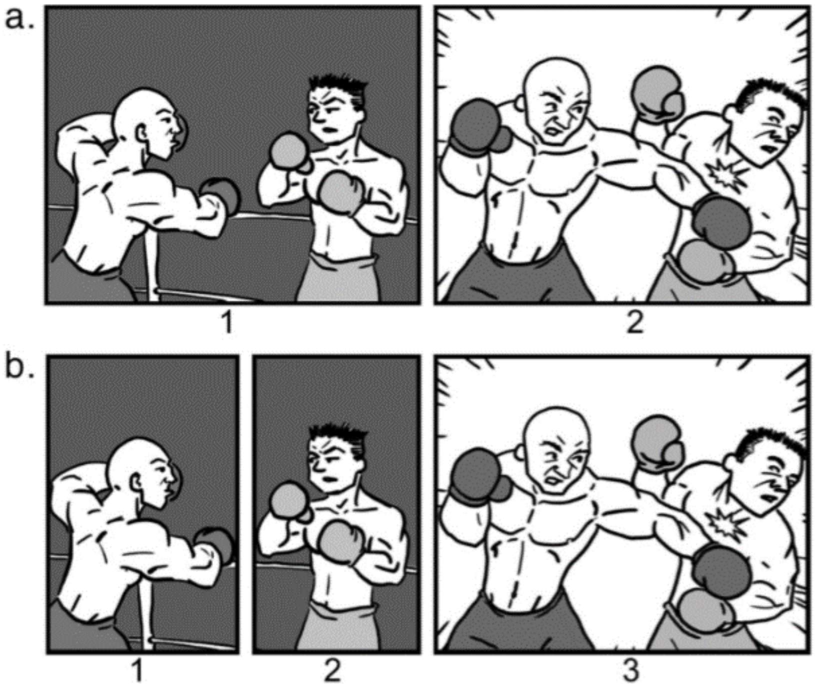


Figure. 4 Variation in how panels provide a ‘‘window’’ on a visual narrative scene (Cohn, 2015)

Group (b) in Figure 4 divides Panel 1 of Group (a) into Panels 1 and 2. Without panel 3 of comic (b), the user probably would not gain the same situational awareness as in panel 1 of comic (a). Only when panel 3 in group (b) is in the same position as panel 1 in group (a), Panels 1 and 2 in group (b) intend to transmit the same message as panel 1 in group (a). Therefore, for sequences containing conjunctions, readers cannot accurately understand what they want to express without adding previous or subsequent panels. For example, in panel 1 and panel 2 in group (b), if the creators do not list panel 3, then the reader will interpret that there are two competitors listed or that the competitor on the left is preparing to hit the person on the right. But when the creators list the third panel, the reader can directly understand that the person on the left is about to hit the person on the right. Therefore, additional semantic reasoning aids the reader in understanding correctly (Cohn, 2015). Furthermore, adding more panels per page facilitates the interpretation of picture data (Foulsham et al., 2016).

In comics, creators use lines to depict the path of moving things. Motion lines, often referred to as speed lines, appear frequently in graphical depictions (Cohn, 2015). Readers find it easier to perceive or recall the direction of items than their movement (Burr and Ross, 2002; Kawabe and Miura, 2006; Kawabe et al., 2007; Kim and Francis, 1998). Thus, the motion line is a classic representation of the visual system-based core component of human perception. In addition, drawing an object's motion lines is more efficient than sketching a line of motion, background lines, or lines flowing in the wrong direction (Cohn, 2015) Compared to images without motion lines, the motion line improves understanding and recall of the depicted event. Abnormal lines make it more challenging for the viewer to interpret the scene than if there were no motion lines (Cohn, 2015). Therefore, the precise direction of motion can clarify relationships between entities that might otherwise remain ambiguous (Brooks, 1977). As a visual language, motion lines can convey several meanings. Similar to how different languages contain words with the same conceptual meaning, different visual languages map various graphical representations to the same conceptualisation path. As a result, in the construction of comics, creators can attach motion lines to moving objects to indicate the path of the action (as shown in Figure 5).

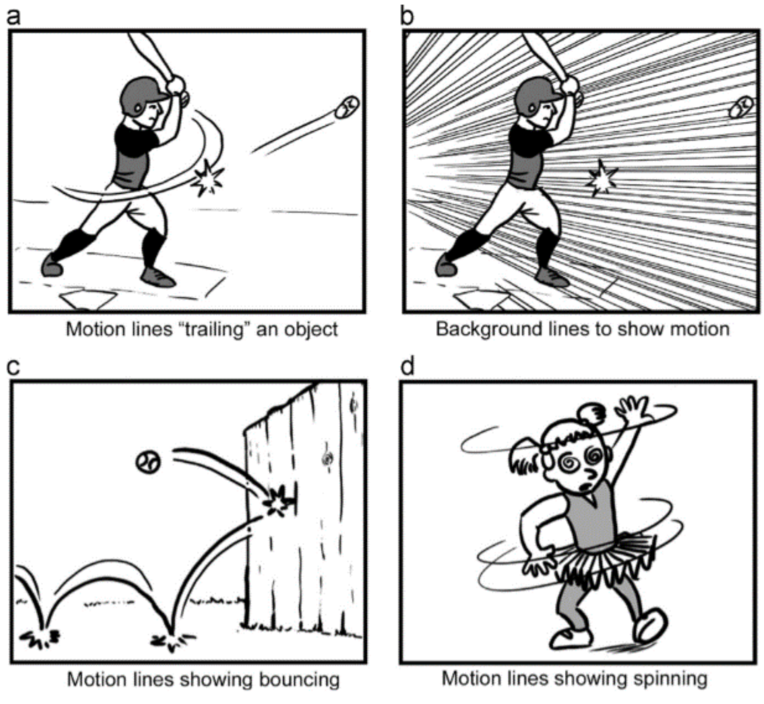


Figure.5 Different uses of lines to depict motion (Cohn, 2015)

In addition to the layout of panels affecting reading strategy, eye-tracking can also reveal how content influences the 'entry point' to the material. From studies of reading newspapers, we know that the size and color of information can determine where people look first (Holsanova et al., 2006). Such entry points may lead to longer fixation time, with the reader developing inferences to guide their reading. These inferences will, in turn, depend on the 'task' (Ballard and Hayhoe, 2009). Reading a comic involves following the plot, enjoying the 'action', interpreting the characters' motivation, and remembering details for later recall (Mikkonen and Lautenbacher, 2019; Cohn, 2013b; Cohn, 2020).

Similar to the arrangement of words in a text, the spatial arrangement of comics often affects the reading order of the reader. For most people, the panels of the comic still follow the text order, the "Z path" (Cohn, 2013a; Brewer and Lichtenstein). But in most comics, the choice of layout tends to encourage readers to deviate from the default reading mode Z path (Mikkonen and Lautenbacher, 2019). In the eye-tracker experiment (Mikkonen and Lautenbacher, 2019), for grid layout, the highest number of participants read using the Z path, regardless of how the panels of the comics changed (Mikkonen and Lautenbacher, 2019). Participants will deviate from the Z path under the layout where the comic panels are arranged vertically. If the creators add blocking factors to the comic, participants choose the Z-path much less frequently compared to the grid layout without blocking. However, regardless of the layout, the preferred path for readers to read comics is still the Z path. The only factor affecting the frequency of the Z path is the participants' background reading habits (Cohn, 2013a).

In conclusion, sequential graphic comics can help participants comprehend the plot quickly and effectively. Therefore, the design of the comic layout will significantly impact the participants' ability to read the visuals. In other words, sequential comics with a tight logical sequence can greatly reduce the time required for readers to comprehend the plot (Cohn, 2010; Cohn, 2019).

In the study of visual language, creators can communicate images through metaphor, metonymy, or a combination of the two to convey more complex conceptions of thought. For instance, the imprint of a lipstick kiss signifies love, and Einstein's haircut signifies intelligence (Cohn, 2010). Moreover, comic book design adheres to the notion of simplicity. Comics do not require extremely complicated graphics because they translate complex information into easy-to-understand images. In comics, simple lines and distinct characters help the reader comprehend the comic's meaning (McDermott et al., 2018).

In many ways, the ability of pictures to convey information transcends many mediums. Whether as ancient cave paintings or on today's electronic devices, pictures are an important part of human experience and access to information (Russell,2011). Levie and Lentz (1980) found that there are five main functions in visual reading that differ from textual reading. These five functions are:

• Presentation: Visually repeats the content of the text or substantially overlaps the text.

• Organisation: Visually enhance the coherence of the text.

• Explain: Visuals provide readers with more specific information.

• Transformations: Visuals target key information in the text and are recorded in a more memorable form.

• Decoration: Visual objects are used for their aesthetic properties or to stimulate the reader's interest in the text. In a meta-analysis on visual effects. Researchers (Levin, 1987) found that all functions except decorative ones contributed to memory. In order of importance, these functions are transformation, interpretation, organisation, and presentation(Jun Liu, 2004).

Providing simple datasets (such as raw data, viz sets, and data stories in video and news article formats) can significantly help participants understand the data (Wang, 2019). Comic books merge text and images. Their interaction effectively conveys information and encourages reader engagement (Hosler and Boomer, 2011). When text and images combine, they improve reading performance and memory compared to text without illustrations. In the experiment by Holser and Boomer (2011), the results confirmed that pictorial information in comics more effectively captures the attention of students. The experimental results show that, compared to traditional books, comic books do not diminish students' ability to gather information (Hosler and Boomer, 2011). Comics aid students in understanding and acquiring information in the form of stories in their education. Simultaneously, compared to text information comics, it provides visual support, which can help non-language readers visually understand the information that texts and pictures intend to convey. Additionally, pictures can help readers make contextual associations and deepen their memory of the information (Ranker, 2007).

While the previous sections have examined how people get meaning from texts and comics, we can consider the knowledge required to make sense of stories as a form of story grammar, which we will discuss in the next section.

1.4 Story Grammars

When reading a story, we consider characters, time, place, events, and outcomes. As already established, users read sports reports with an event-centric perspective. By analysing a story using story grammar, the reader can identify important events. In our system, we employ the concept of story grammar to identify characters, events, and their relationships. Our thinking builds on the concepts of Formalist literary analysis(Pomorska, 1985), which asserts that basic plot structures exist. For the Russian Formalists, the plots pertained to folk tales, but a similar concept has proposed seven basic plots for Hollywood movies (Booker, 2004). This implies that summarising a sports report could involve covering the basic grammar of the 'plot' of that particular sport. In other words, if the text presents a soccer game report, then the generic grammar of soccer games can frame the identification of characters, motivations and key events that require extraction. These elements can map onto the output of the text extraction (figure 11).

When reading a novel, the reader takes an interest in the characters, time, place, events, and the resolution of the events. We can enhance understanding and minimise the need for interference by structuring the narrative in line with story grammars (Black and Wilensky, 1979; Brewer and Lichtenstein, 1980; Robert, Behavioral Wilensky and Sciences, 1983). However, when people read a story, they form a gist (as previously noted) which comprises the words they read and the inferences they use to fill in the gaps. For instance, a story might contain sentences such as, 'Joe's ice cream fell to the floor. Joe began to cry. Even from such minimal information, the reader can assume Joe's age and his reason for crying - neither of which the sentences state but which result from the reader's inference based on previous experience and knowledge.

Since readers focus mostly on the event, it is important to breakdown the “event” into several components. The grammar of Rumelhart (1975) is a fundamental, straightforward tale grammar comprising the fundamental elements.:

Rule 1: story = setting + episode
Rule 2: setting = (state)*
Rule 3: episode = event + reaction

This structure can be visualised as follows:

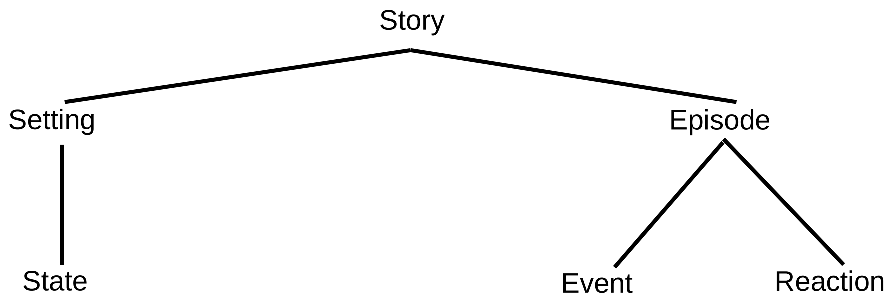


Figure 6. The structure of Rumelhart's grammar (Rumelhart, 1975; Wilensky, 1983)

Thorndyke expanded Rumelhart's Grammar by creating a grammar that includes a more thorough explanation of the tale grammar. The organisation of Thorndyke's grammar is (Black and Wilensky, 1979; Thorndyke, 1977):

Rule 1: Story = Setting + Theme + Plot + Resolution

Rule 2: Setting = Characters + Location + Time

Rule 3: Theme = (Event)* + Goal

Rule 4: Plot = Episode

Rule 5: Episode = Sub goal + Attempt + Outcome

Rule 6: Attempt = Event* or Episode

The composition is as shown in figure 7.

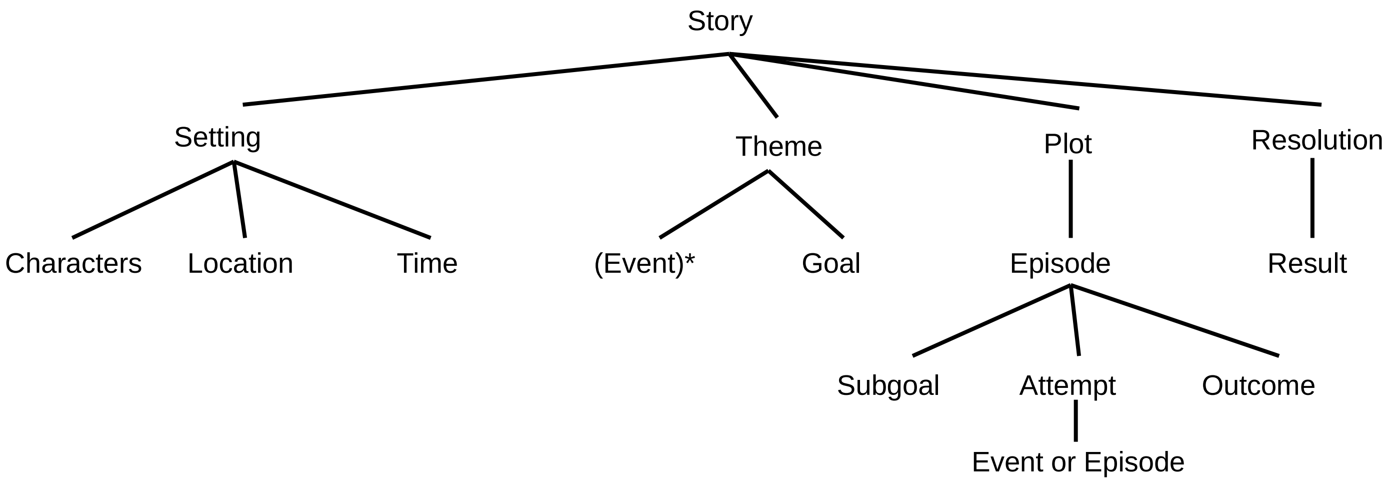


Figure 7. The structure of Thorndyke's grammar (Thorndyke, 1977)

The essential structure of a narrative grammar is the same for all stories, despite significant changes between the story grammar of figures 6 and 7. Therefore, the reader can easily identify the key events in the story by utilising a story grammar to analyse it.

As part of this PhD research, we used the concept of story grammar to help summarise the text. This will be explained in detail in chapter 2. However, it should be obvious that the approach to text summarisation employed makes use of a story grammar that defines typical events associated with soccer games. In the next section, approaches to text summarisation are explored.

1.4.1 Events in Sports Report

Generally speaking, sports fans expect to be able to obtain the game situation of the team they follow from the scene, video and news reports. Moreover, sports competitions are the most entertaining competitions, paying attention to and discussing the relevant details of sports competitions can bring great pleasure to people. Therefore, many readers will quickly obtain content related to the game by reading the game-related reports. In the registration of sports competition, it generally contains several types of events:

- Game Results: This is the most common and most talked about sports coverage content.
- Game process: This is a report on the behavior of athletes during the game and some emergencies that occurred during the game. For example, a long-range shot in a basketball game, a volley in a football game, and so on.
- Player/Athlete Information: Personal information about relevant players in the game, such as injury status, previous personal performance, etc.
- Event analysis, etc.: In-depth analysis and interpretation of completed events.

If we consider a sports game, such as a soccer game, there will be a finite set of events that could typically be associated with the story of this game. In soccer, there will be two teams competing for goals, their competition will take place according to the rules of the game (overseen by match officials such as referees), and a crowd of supporters will watch their performance.

We surveyed 40 football match reports (print newspapers and online). This was performed to identify 'events' in these reports. Through the event statistics of football match reports, we present the results in the form of histograms (Figure 8). Through the histogram, we can find that the events with the highest frequency in the current game are: shooting, accident, injury, foul, etc. At the same time, these are the types of events that readers are most concerned about. This will help us to summarize and extract the event registration in the later stage.

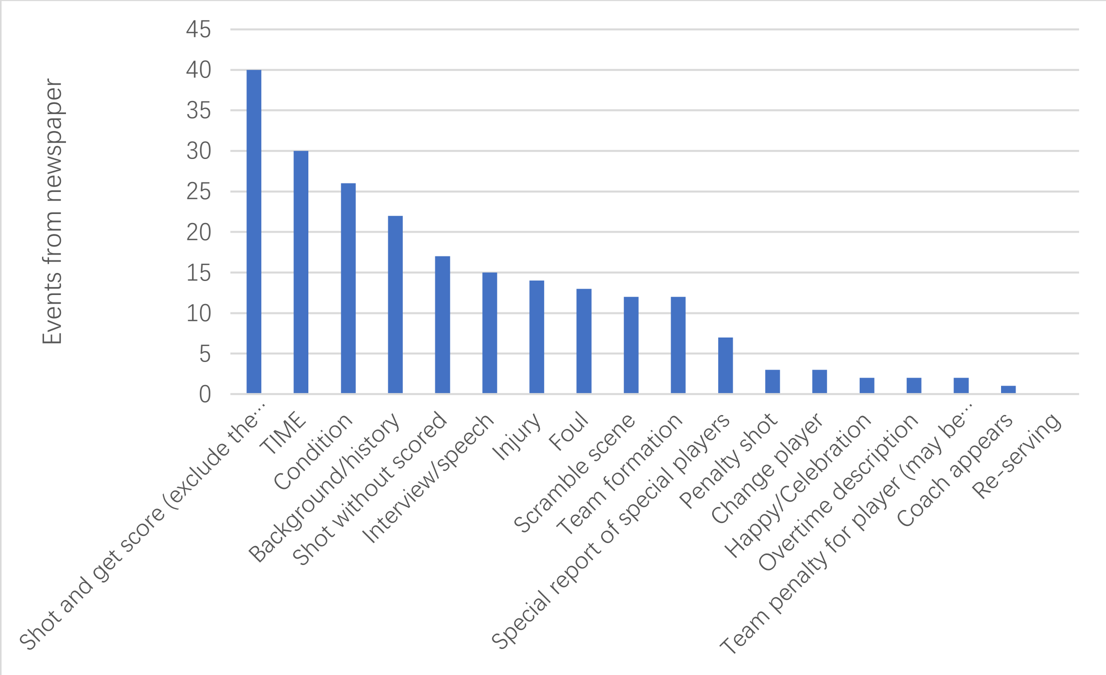


Figure. 8. The statistics histogram from the newspaper

1.5 The Purpose of this Thesis

The existing text-to-image visualization primarily focused on data information visualization, which outputted chart-type pictures. Details of related work will be described in the next section. GAN-based text visualization also suffered from vagueness and imprecision. Current deep learning models, such as GAN models, have yet to achieve accurate visualization of multiple events. Therefore, to address the multichannel visualization of combined events, we aim to design and build a new text visualization system based on story grammar. We begin by understanding how humans extract necessary information from textual data through the analysis of human reading behaviour on text. This understanding helps to extract practically meaningful words through linguistic grammar. Studying comic reading behaviour also assists us in constructing comic's essential content, defining the comic's content structure, and arranging multichannel layouts.

The core of our design lies in the story grammar. We will integrate story grammar with linguistic grammar to extract key words suitable for story grammar through natural language analysis. This integration will aid in constructing special dictionaries based on story grammars. Hence, the system can filter and extract sentences containing event information based on the story's grammar and assemble them into summary text. We will utilize the extracted summary text for visualization. Our visual system for summarizing the text also constructs the story grammatical structure. We will utilize the visualization system to visualize multiple events in the summary text and align them in multichannels along a timeline. In the end, we can conduct experiments to study the reading behavior of long and short texts and the corresponding comic material, and understanding capabilities under different materials.

1.5.1 Contribution of this Thesis

Comics exhibit an ability to attract participants' interest in a topic and aid comprehension compared to text alone. Thus, converting text information into comic information greatly assists non-professionals and individuals with language barriers. Our research goal also involves designing a text visualization system and evaluating the comprehensibility of different types of materials.

In this study, we diverge from previous text summarization methods such as statistical methods. We create a text summarization system based on story grammar, merging linguistic grammar with story grammar. The system can summarize the original text based on the story's grammatical structure and a specialized dictionary. This process retains the original text's crucial event information and semantic coherence while largely addressing information redundancy issues. Ultimately, we successfully convert the events contained in the summary text into pictures based on the story grammar and relevant dictionaries. We then assemble them into a complete cartoon according to the timeline.

This approach varies from previous data visualization and text visualization for GNN models. This method not only visualizes the essential events contained in the text information and assembles them into cartoons that can express the original text but also reduces manual labelling work and model training and fine-tuning. Consequently, we achieve accurate text visualization with significantly reduced labour costs.

Through the analysis of audience reading experiments, we found that comics are more helpful to improve the audience's understanding and interest than text. And through the eye tracking experiment, we found that the participants received the least interference when reading the comic material, and could quickly focus on the key information contained in the event.

2 Review of Related Work

2.1 Summarising Text

The objective of automatic text summarisation is to minimise the quantity of text while preserving its essential meaning in order to lessen the text's complexity(Gambhir and Gupta,2017; Kucher and Kerren,2015). This could mean, for instance, that the summary should be more understandable than the original. Using individuals to generate abstracts could result in substantial use of human resources (Yang et al.,2012). Consequently, computerised text analysis techniques are crucial for large amounts of text. It can considerably minimise the consumption of human resources and the rate of knowledge loss, and help users cut retrieval time. It effectively addresses the issue of information overload (Gambhir and Gupta, 2017). Before discussing approaches to summarisation, a brief discussion of linguistic grammar might be instructive, as it informs the methods used to describe and extract textual elements.

2.1.1 Linguistic Grammar

Around ten basic grammatical components can be present in every sentence: nouns, verbs, adjectives, adverbs, pronouns, numerals, articles, conjunctions, prepositions, and interjections(Suhrob and Vasila, 2022; Kovbasko,2020). In the text, many speech acts have different functions (Postal, 1964; Rosenbaum, 1965). In general, the basic building blocks of a sentence are nouns and verbs. The phrase's standard form is:

Sentence = Noun (subject) + Verb (predicate) + Noun (object)

As supplemental vocabulary, adjectives, adverbs, and other elements of speech will be employed to alter current phrases (Rosenbaum, 1965). For example, red apples and green apples are different semantics. The basic function of conjunctions is to establish a fair, logical relationship between two phrases and to reinforce the expression of the sentence. Figure 9 depicts the basic structure of linguistic Grammar.

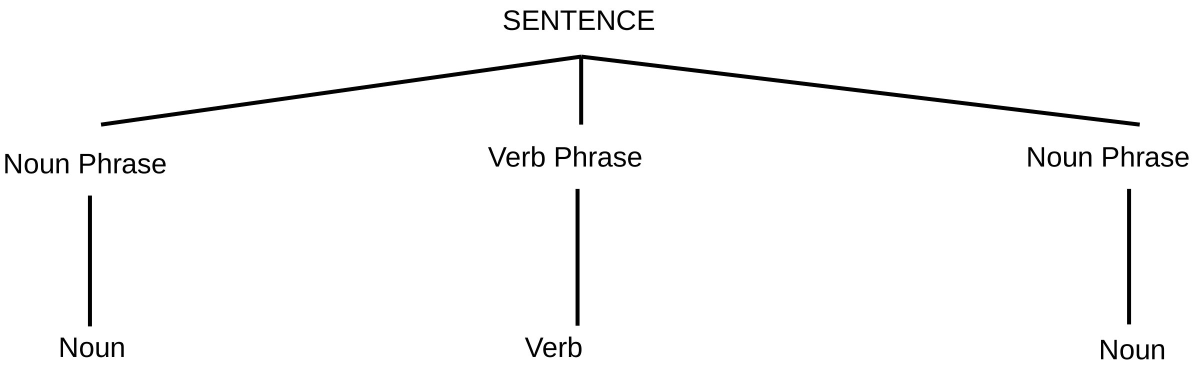


Figure 9. The basic structure of linguistic grammar (Postal, 1964)

In actual text applications, however, the phrase structure is more complex. Frequently, we use adjectives, adverbs, conjunctions, and other lexical items to alter sentences, so enhancing their complexity and expressiveness. According to Rosenbaum (1965), a more complex linguistic grammar can be stated as follows:

Sentence = Adjective / Article + Noun + Verb + Adverb + Adjective / Article + Noun

Sentence 3 = Sentence 1 + Conjunction + Sentence 2

Therefore, a complex linguistic grammar structure can be expressed as (Postal, 1964; Rosenbaum, 1965) :

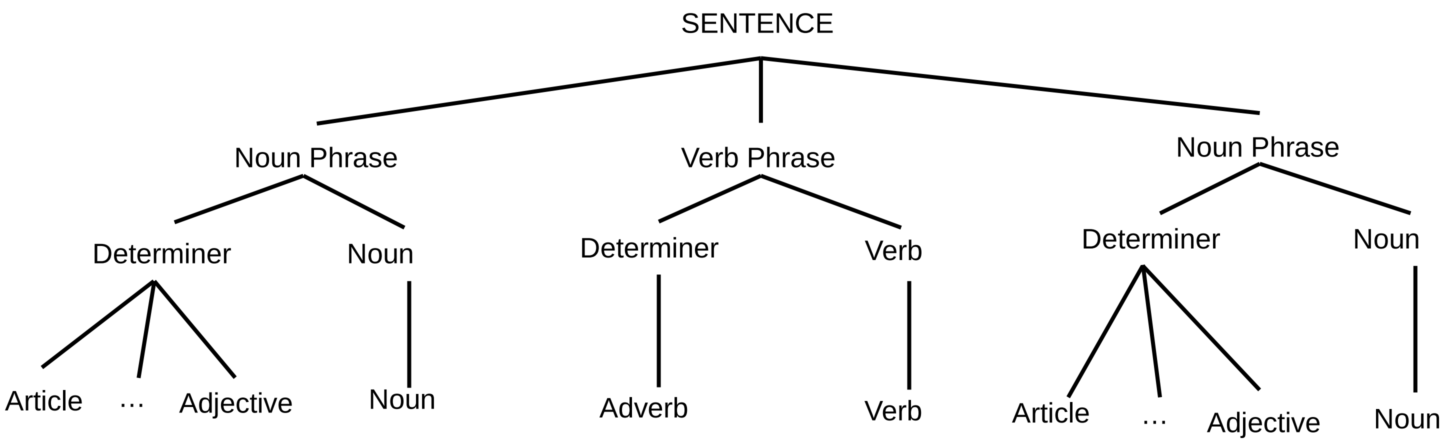


Figure 10. The complex structure of linguistic grammar

In essence, a language record is a text with a complete storyline, composed of many sentences (Rosenbaum, 1965). Beyond the basic form of linguistic grammar previously discussed, linguistic grammar also encompasses more complex scenarios, like inverted sentences and interrogative sentences. Although variations in sentence patterns add to the complexity of linguistic grammar, all grammar bases itself on the meaning and part of speech of words. Additionally, subjects, predicates, and objects form integral parts of every language grammar.

Transferring linguistic grammar onto the narrative grammar covered in the previous section is straightforward. The blue components of Figure 11 make up a story grammar. This shows that the noun and verb sections of the grammar correspond to different parts of the narrative grammar. Hence, narrative grammar is more advanced than linguistic grammar. Story grammar is the macrocosm of linguistic grammar, while linguistic grammar is the microcosm of story grammar. To elaborate, story grammar refers to the set of rules and structures that guide the organization and progression of a story, such as character roles, setting, conflict, resolution, etc. It provides a framework for constructing and understanding narratives. Conversely, language grammar is the set of structural rules that dictate the composition of words, phrases, and clauses in any given language. It defines how we construct and understand sentences and phrases. In a way, we can say that story grammar and language grammar are related - language grammar builds the sentences and paragraphs that constitute a story, while story grammar provides the overall structure and flow of a narrative. We can see each sentence in a story as an example of the grammar of the language at work, and the overall structure of the story is an application of the grammar of the story. Concurrently, in linguistics, verbs usually express an action or state, while nouns typically refer to a person, thing, place, or concept. In sentence structure, verbs and nouns form a predicate and a subject or object respectively, forming the core components of a basic sentence. However, in the structure of a story, we often abstractly understand "verbs" and "nouns" as part of elements. In storytelling, we can understand "verbs" as actions or events in a story, while we can understand "nouns" as characters, things, or settings in a story. By merging narrative and linguistic grammar, we reduce the complexity of the text while preserving the meaning. This aids the algorithm in selecting the most relevant information during natural language recognition.

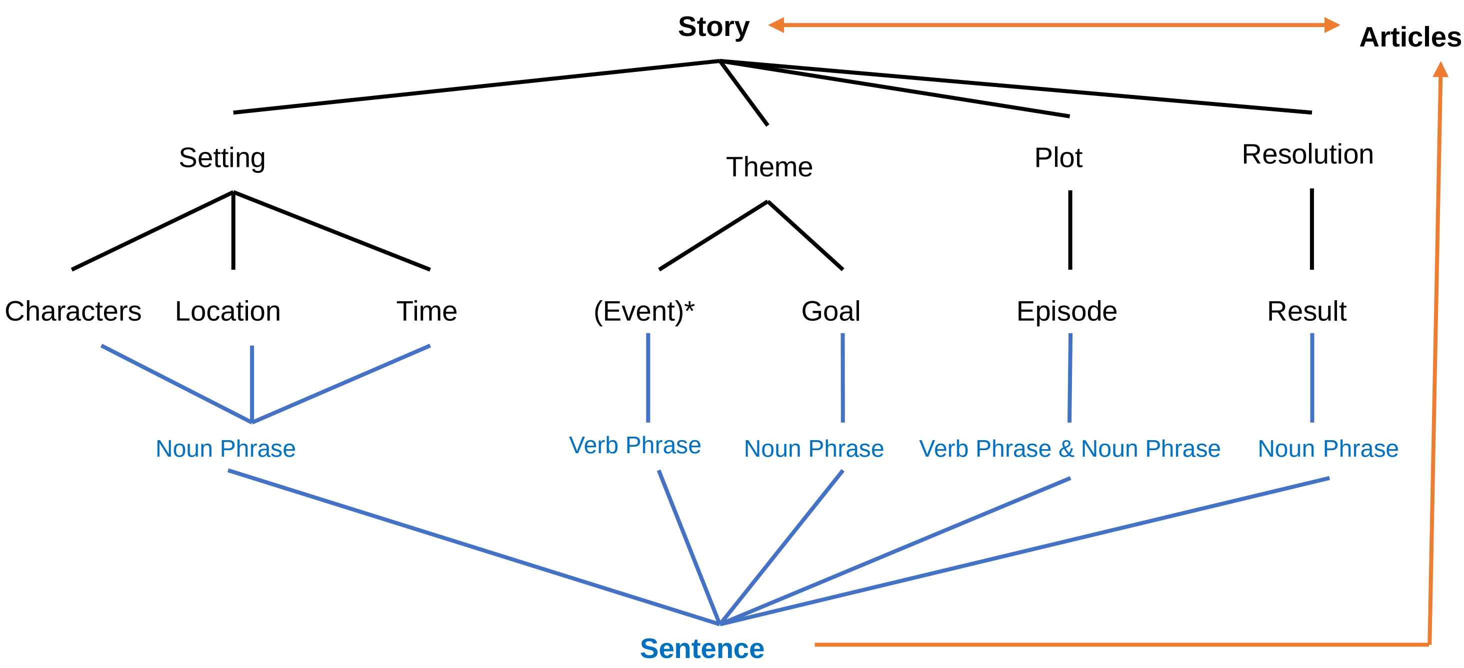


Figure 11. The relationship between the two grammars

2.1.2 Approaches to Text Summarisation

In the simplest terms, text summarization involves either (i) extraction, which identifies and removes prominent terms from a document, or (ii) abstraction, which rewrites the document in a shortened form while conveying its meaning. The extraction approach does not promise a 'meaningful' outcome or provide a smooth report; instead, it delivers a set of statements in the order they appear in the document without explicitly linking them. The abstraction approach is more challenging and requires deeper processing to produce a meaningful and smooth output for the reader. Our approach applies extraction (with Spacy tools in Python) and enhances it with two additions specifically for our purposes: (a) a 'story grammar' of soccer games and (b) a 'dictionary' of terms used in sports reports for soccer games. These allow the reconstruction of terms and entities extracted from the text to support the creation of comics.

The most common method for text summarisation uses statistics to extract words, phrases, and sentences from a text and integrate them into a new text while maintaining the original text's logical order (Berger and Mittal,2000;Carbonell and Goldstein,1998;Nomoto and Matsumoto,2001;Strzalkowski et al.,1998;Zechner,1996). Typically, these strategies look for statistical indicators to establish relevance (Fattah and Ren,2009;Gambhir and Gupta,2017), such as keywords (based on frequency counts), sentence centrality (i.e., resemblance to other phrases), and proper nouns (named entities) in sentences. These systems can build summaries by selecting the most relevant sentences and combining them in the original text's textual sequence (Zechner, 1996).

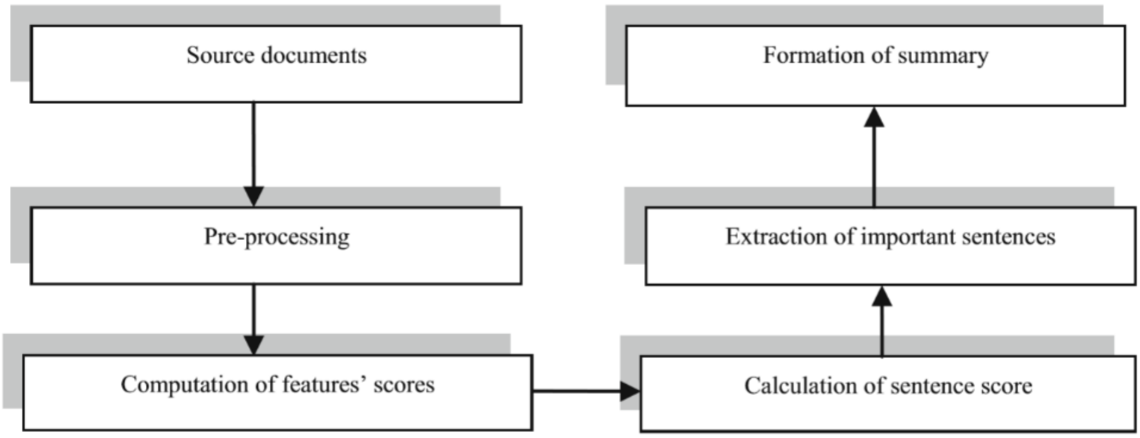


Figure 12. Block diagram of automatic extractive text summarisation system by using statistical techniques (Gambhir and Gupta,2017)

Due to low correlation and duplication, however, the results gained using statistical methods are limited. For instance, the extracted sentences or words may be very redundant and contain partially or entirely redundant information (Carbonell and Goldstein, 1998). Consequently, eliminating redundancy is an essential step. Using Maximum Marginal Relevance, statistically-based text analysis systems eliminate redundancy and maintain query relevance (MMR). This strategy is investigated to eliminate phrase repetition greatly (Carbonell and Goldstein, 1998). Maximum boundary correlation algorithm is another name for the MMR algorithm. This technique was originally intended to rank documents based on their closeness to the query text (Carbonell and Goldstein, 1998).

$$MMR(Q,C,R) = \underset{D_i \in R \setminus S}{argmax}[\lambda sim_1(Q,d_i) - (1-\lambda)\max_{D_j \in S}(sim_2(d_i,d_j))]$$

D_i: a document in collection C;

Q: query;

R: A collection of related documents in C;

S: current result set;

λ: Adjustable hyperparameter, the larger the λ, the higher the accuracy; the smaller the λ, the lower the accuracy;

However, due to the limitations of MMR, the system cannot minimise redundancy effectively. The formula splits into two parts: the left-hand side represents the similarity between the candidate sentences and the query, while the right-hand side represents the maximum similarity between the candidate sentences and the set of all selected sentences, with a negative sign. This means that lower similarity between the final candidate sentences is better (Carbonell and Goldstein, 1998).

Traditional summarisation techniques primarily determine sentence characteristics based on statistical or linguistic data, allowing these characteristics to extract phrases or paragraphs that hold an important position in the document. Therefore, minimising redundancy is a necessary step. Statistically-based text analysis systems use the Weighted Average Algorithm (WAA) or Maximum Marginal Relevance to decrease redundancy and maintain query relevance (MMR). A main drawback of the traditional strategy is that it overlooks the sentence's semantic relevance (Zhang and Li, 2009). This significantly minimises sentence repetition (Carbonell and Goldstein, 1998). Zhang and Li (2009) suggested a strategy for the automatic generation of text summaries based on the recognition of the semantic relationship between sentences, which they termed "sentence clustering." This method enhances the relevance of extracted sentences compared to traditional text analysis techniques and reduces the incidence of redundant information. This technique boosts the accuracy of the automated abstraction system, as demonstrated in Table 2 by Zhang and Li's (2009) comparison with other methods using the document dataset DUC2003.

TABLE 2 THE VALUES OF EVALUATION METRICS FOR SUMMARISATION METHODS (DUC2003 DATASET) (Zhang and Li, 2009)

Methods	ROUGE-1	ROUGE-2	F1-Measure
MMR	0.34813	0.07917	0.43245
WAA	0.38023	0.09121	0.45334
Sentences clustering	0.43512	0.10142	0.47576

Note that in table 2, ROUGE represents a collection of measures for evaluating text summarisation. ROUGE-1 assesses the overlap between each word in the generated summary and the reference (gold standard) summary, whereas ROUGE-2 examines the overlap of bigrams (i.e., two-word sequences) in the created and reference summaries. ROUGE did not specify if the summaries reflect the original source's meaning or if they are easy to read.

NLP helps understand and modify natural texts or voice (Chowdhury, 2003). For instance, Named Entity Recognition (NER) is a method that extracts entities from a text. NER identifies and highlights the names of people, places, and organisations in the text (Ratinov and Roth, 2009). Often, these entity words or phrases represent the most significant and representative parameters of the text. Extracting and analysing named entities allows users to identify the theme of the text more quickly. At the same time, identifying named things (such as names of people, places, etc.) can help the visualization system directly recognize object information when visualizing text, thus saving workload and improving accuracy.

In 2001, the University of Pennsylvania's Computational Linguistics Program developed the Natural Language Toolkit (NLTK) in partnership (Loper and Bird, 2002). NLTK analyses text information based on factors such as part of speech, sentence relationship, and context. SpaCy is another tool that specializes in large-scale information extraction activities (Vasiliev,2020). SpaCy performs not only the operations of NLTK, but also provides a more comprehensive and production-oriented language analysis. Additionally, SpaCy can generate more precise results by training models on large amounts of data, thus ensuring the accuracy of processing results. Figure 13 shows a screenshot of the chart parsing demonstration.

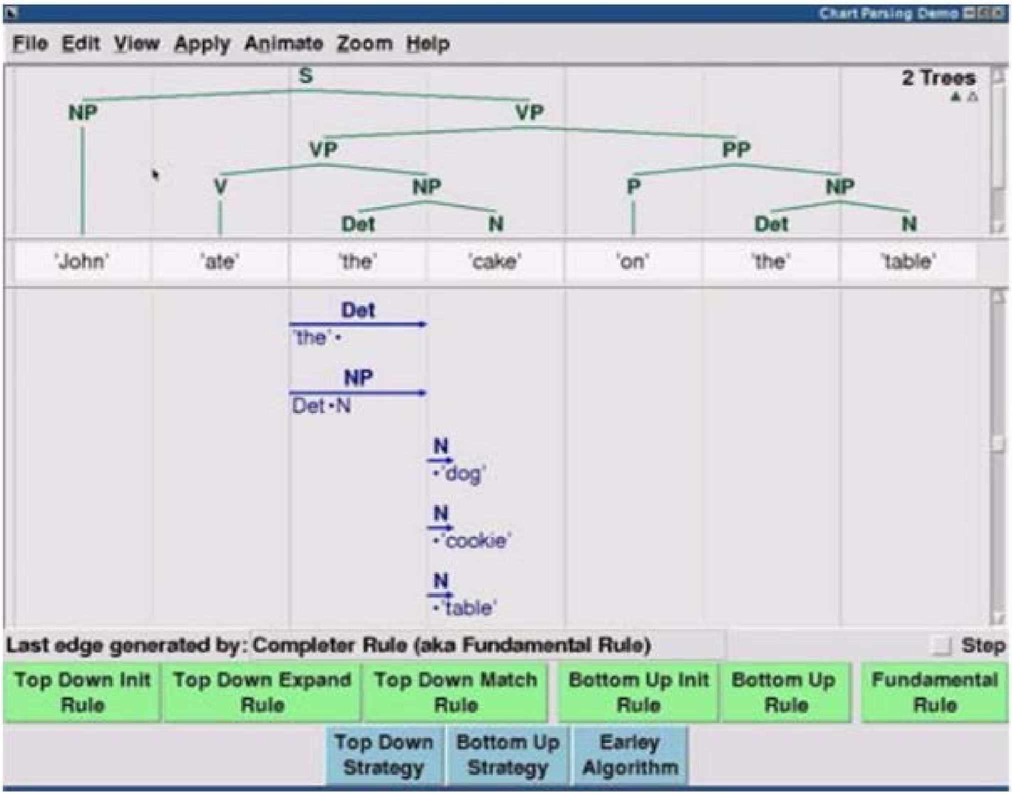


Figure 13. Interactive Chart Parsing Demonstration (Loper and Bird, 2002)

SpaCy is a free, open-source Python package for advanced Natural Language Processing. SpaCy excels in extracting information on a massive scale. Developers wrote it from scratch in Cython, managing memory strictly. Independent research (Saabith, Vinothraj and Fareez, 2020) confirms SpaCy as the world's fastest protocol. SpaCy uses Cython to boost the performance of compatible modules. Its language analysis performance outdoes Python NLTK. SpaCy doesn't only perform NLTK's functions, but also enables a deeper examination of language. SpaCy has greater capabilities than NLTK. This is because Spacy supports functions such as deep learning framework and word vector, and it offers higher speed and accuracy in text processing than NLTK. Training the model will improve the accuracy of SpaCy's output. The grammatical function of each word in a sentence plays a crucial role in defining its meaning. Therefore, acquiring parts of speech is essential to the process of linguistic Grammar-based analysis. By analysing text with tools for natural language analysis, we can determine the part of speech and significance of each vocabulary word in the current text. Analysis also helps identify the dependencies between vocabularies, assisting the system in locating dependencies between modifiers and objects.

Both of the aforementioned NLP-based technologies use the Python programming language. Python is an interpreted, object-oriented language that allows interactive exploration (Loper and Bird, 2002). Python enables data and method encapsulation and reuse. It also has a comprehensive standard library that includes graphical programming and digital processing facilities. Python facilitates the development of interactive algorithm implementations through these standard libraries (Loper, 2004; Van Rossum and Drake, 2003; VanRossum and Drake, 2010).

2.2 Previous research on visualising text

Data-driven storytelling, utilised in various fields such as astronomy (Meuschke et al., 2022; Ma et al., 2011), medicine (Meuschke et al., 2022), and sports (Perin et al, 2013), emphasises communication via data visualisation. The objective of data-driven storytelling is to communicate effectively using visualisation tools based on or containing specific data. For complicated texts, users often find it difficult to easily discover and collect critical information from large text collections or locate key information in vast text sets. This complexity makes understanding and parsing the text more difficult and time-consuming for users (Vi et al., 2006). However, by using a data visualisation system, transforming a summary into a more understandable graphic system can improve users' comprehension of the article's subject matter (Shixia Liu et al., 2009; Segel and Heer, 2010; Wise et al., 1995).

Data visualisation through charts can convey information more efficiently and engage the reader's attention (Cui et al., 2018). Infographics, which visualise data, mainly convey complex information through data visualization (Bateman et al., 2010; Haroz, Kosara and Franconeri, 2015; Harrison, Reinecke and Chang, 2015). The design of infographics traditionally poses a difficult and highly creative task. Usually, professional designers, assisted by computers, create this work, and it's difficult to complete it by computer alone. Therefore, non-professionals, who are not only unfamiliar with professional design tools but also lack in design theory, often struggle (Cui et al., 2018). As a result, there has been extensive research on the design and development of graph information (Borkin et al., 2013; Hullman, Adar and Shah, 2011; Kim et al., 2017; Satyanarayan and Heer, 2014; Skau and Kosara, 2017; Wang et al., 2018). These research efforts aim to facilitate the creation of data-driven infographics. Although these research works and related tools strive to balance ease of use and functionality, these tools are not user-friendly for casual users (Cui et al., 2018).

Cui et al., (2018) use natural language processing models to create a novel system for automatically generating infographics from text sentences. They acknowledge that infographics have many types of expressions, and it's impossible to cover the entire space on a single page. Therefore, they focus the design on a relatively and isolated information space and build a proof-of-concept system (Cui et al., 2018) for it. The system has a supervised CNN+CRF model. The working principle is to convert the text into a tag sequence through preprocessing, and then extract features for each tag to achieve tag characterization. They extract word embedding features (Mikolov et al., 2013), syntactic features (e.g. uppercase/lowercase, punctuation, part-of-speech tags), and Brownian clustering features (Brown et al., 1992) here. They then aggregate these features into a large vector that can represent tokens. Finally, they use the CNN+CRF model to process and analyze these tokens. Cui et al. (2018) trained the model on 800 manually annotated sentences. Figure 14 shows the output of Text-to-Viz. It demonstrates that different templates generate different results. However, the system still has many limitations. For example, complex and redundant long texts generate incomprehensible charts, and currently, it can only deal with fewer and specific types of information sets. Moreover, the expressive power of icons does not have human creativity (Cui et al., 2018).

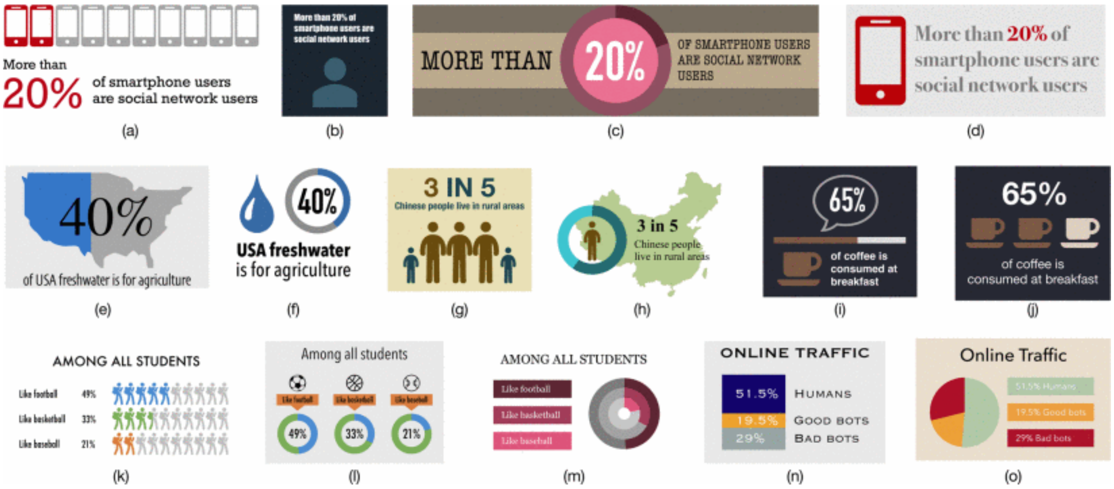


Figure 14. Examples created by Text-to-Viz (Cui et al., 2018)

The transformation of data into charts can effectively assist users in creating infographics based on data sets. However, such systems struggle to visualize text information containing events effectively, as the expression of events usually exhibits complexity and redundancy, with more representations through words rather than data.

Liu et al. (2009) proposed a system for time-based visual summarisation of text. The system primarily employs the Latent Dirichlet Allocation (LDA) model to ascertain the probability distribution of a document's subjects. Then, they pair it with the interactive visual text analysis program TIARA (Text Insight via Automated, Responsive Analysis) to deliver visual picture data with interactive features and keyword clouds (Shixia Liu et al., 2009; Wei et al., 2010). They show the display in Figure 15. However, 80% of users reported that the resulting visual summary seemed unclear under specific conditions. Consequently, users committed distinct errors in the comprehension process (Shixia Liu et al., 2009). It's safe to say that this type of visualisation doesn't assist in overcoming linguistic difficulties. Moreover, as shown in Figure 15, the system can only visualize crucial textual information and not the entire narrative. Therefore, the application is restricted and can't find wide utilization.

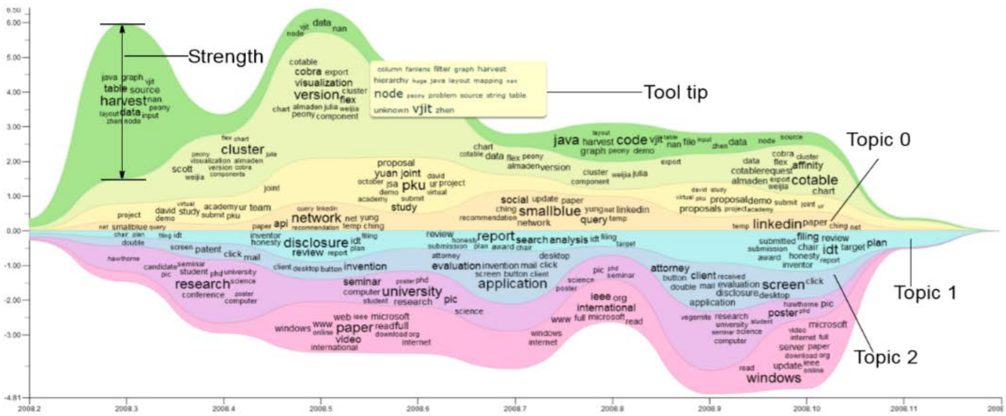


Figure 15. The example of visual abstracts created by TIARA (Liu et al., 2009)

Regarding the visual analysis of sports, it primarily concerns the analysis of trajectory and spatio-temporal information of the game (Gudmundsson and Horton, 2017; Stein et al., 2017). Among all kinds of social competitions, team sports serve as extremely important recreational activities. Each team employs various methods to analyse the opponent's performance to enhance their win rate. The objectives of team sports analysis often encompass identifying weaknesses of the opposing team, or assessing the performance and potential for improvement of a coached team (Stein et al., 2017). People have performed manual notation of events in soccer matches using manual notation methods since the 1950s (Reep and Benjamin, 1968). But the experimental results of Franks and Miller (1986) indicate that human recall of content drops as low as 42%. Many existing systems can capture spatiotemporal data during games, simplifying the process for researchers to analyse and process various behavioral trajectories during games (Gudmundsson and Horton, 2017). Gudmundsson and Horton (2017) analysed game spatio-temporal data, such as the intensity map of events during the game (including the division of personnel distribution areas); the transfer network between players, and so on. They pointed out that the most commonly used visual aggregation of spatiotemporal data in the game is a heat map. Heatmaps can visualize various data simply and intuitively. For instance, by visualizing the distribution and range of basketball shooters, it's possible to identify the best shooters (Goldsberry, 2012). The SoccerStories system proposed by Perin, Vuillemot and Fekete, 2013 (2013) serves as a visual interface for football game data. It can draw athletes' trajectories and distribution heat maps based on the spatio-temporal data in the game, which helps analysts to analyse and summarize the game (Perin, Vuillemot and Fekete, 2013).

However, the system remains based on data visualization, and does not convert text-based event information into comprehensible images. We propose a novel method for text visualisation based on story grammar. We will use this with our new technology for automatically summarising text to generate images that reflect specific occurrences. We will provide a detailed description in the Methods section.

2.3 The Visualization of Comics

Segel and Heer (2010) categorized visualizations into seven genres: magazine style, annotated chart, partitioned poster, flow chart, comic strip, slide show, and video. Later, Bach et al. (2018) refined these into three more general types - space-oriented, time-oriented, and comics. Space-oriented visualization typically involves graphics or charts to present as posters, webpages, or leaflets. It emphasises the use of graph elements such as layout organization, 3D models, art designs, and allows users to interact at their own pace. For instance, Meuschke et al (2022) and colleagues employed a series of "storyboard" to communicate numerous diseases with avoidable risk factors. They adapted 3D anatomical models in their storyboard design to illustrate how diseases occur on specific organs or blood vessels. Users interacted with the story by touching the screen. After comparing storyboard visualization to standard web blog visualization with minimum interactive activities, the researchers discovered that participants who engaged with the storyboard remembered more information about the disease. They concluded that when communicating diseases to patients using data-driven storytelling techniques, narrative design, user involvement, and easy interactions should be considered. Therefore, space-oriented visualizations with 3D models and easy interaction demonstrated practicality and are worth exploring further in the medical field.

Another type of visualization is time-oriented. It emphasizes time-sensitive information and presents information in a linear manner with a fixed beginning, middle, and end. Sports studies have adopted this method. Researchers designed SoccerStories (Perin, Vuillemot and Fekete, 2013) to analyze soccer data. Using this visualization, users can see the trajectories of certain player actions in a timely manner. A soccer journalist, with no prior experience working with this new system, favoured this visualization (Perin, Vuillemot and Fekete, 2013). Furthermore, SoccerStories allowed the journalist to explore new insights of a game (Perin, Vuillemot and Fekete, 2013). Although researchers have challenged that the use of time-oriented visualization may not be tailored to individuals since some people may consider the flow too fast and too much (Tversky, 2002), SoccerStories avoided such limitations by adding interactive features (such as timeline navigation) (Perin, Vuillemot and Fekete, 2013).

Comics also serve as a data-driven storytelling method. Comics possess features of both space-oriented and time-oriented visualisation. While the layout of comics is subject to graphic designs, it can tell a story following a certain timeline. Because of the flexible nature of comics, it is challenging to define what data comics are. Bach et al. (2017) proposed four elements of data comics – data visualization, flow, narration, and words-and-pictures. Bach et al. (2018) argued that not many examples combined all four elements when attempting to create data comics, and therefore, they proposed data-comic design patterns to facilitate future relevant research. They refined data-comics patterns into several categories – narrative patterns, temporal patterns, faceting patterns, visual encoding patterns, granular patterns, and spatial patterns. After running workshops with students aged 22 and above to implement data-comics in these patterns, researchers concluded favourable outcomes from students with no prior experiences in data-comics. Students further confirmed its usefulness by giving an average rating of 5.6 on a 7-point Likert scale. Thus, data-comics have shown potential for future research in data visualization and data communication (Bach, 2018).

There are at least two languages for describing comics digitally. They are Comic Book

Markup Language (CBML) and Comics Markup Language (ComicsML) (Alves et al., 2007). Compared with CBML, ComicsML is a simpler descriptive language, providing basic information for each panel of comics, such as speech balloons and related image links. CBML, on the other hand, is a higher-level language. Its aim is to provide methods related to the digitisation of comics to improve the preservation time of comics. CBML allows more advanced information, such as publishers and lists of related characters, somewhat similar to text messages in speech balloons in ComicsML. ComicsML focuses on comics for digital media, while CBML focuses on comics for traditional media. However, both languages are limited to face-to-face languages and cannot adapt to more complex comic layouts (Alves et al., 2007).

As previously mentioned, we can use LDA and TIARA for the visualisation of text data topics, providing users with more intuitive access to relevant topics in the article. One of the limitations is that while this type of visualisation visually displays key information about the text, it may create language barriers for users speaking different languages. Therefore, we can improve this kind of visualisation with the image being visually synthesised, thereby producing different pictures with a certain logical relationship. This series of pictures helps users understand the main content of the text and helps overcome language barriers to the greatest extent.

The field of artificial intelligence also studies synthesizing images from text (Reed et al., 2016). It uses neural networks and machine learning to analyse text input and generate two-dimensional or even three-dimensional visuals. The Generative adversarial networks (GAN) confrontation network is one of the most important algorithmic network in this subject (Reed et al., 2016). Using the technique of deep convolution, the image obtained from the text undergoes continuous recognition and rectification until we obtain the best accurate image in the current training set (Reed et al., 2016). Enhancing the confrontation between **discriminator and generator** and increasing the corpus data can effectively improve the accuracy of image generation. However, the current results cannot fully meet the user's needs for picture clarity and accuracy. Therefore, GAN needs to be trained to continuously obtain the best pictures. However, this method requires a large amount of training data sets, and at the same time cannot visually generate multiple events on the same page at the same time. Although some models can generate fairly realistic images, there are often some unnatural or unphysical aspects, such as unnatural colours, shapes, and proportions.

Therefore, we built new visualization techniques based on soccer match check-ins. As such, the visualization system can generate visualizations of multiple events at the same time and finally output an image covering the complete time axis. At the same time, through the application of an accurate image library, we have largely solved the problem of unclear image expression.

3 Methods

Automated text summarisation technology transforms complex texts into basic summaries, thus reducing the need for human resources. Designers of text summarisation systems aim to prevent information overload while retaining vital content information (Yang et al., 2012; Gambhir and Gupta, 2017). The primary purpose of information visualisation technology is to convert the most streamlined text gathered into image data. This is because the internal image information assists the user's visual understanding of the text information (Shixia Liu et al., 2009; Kucher and Kerren, 2015).

As I outlined in chapter 1, comics serve as a type of visual narrative that uses sequences of images (with or without text) to transmit information. For Eisner, a comic is "an arrangement of pictures or images and words that narrates a story or dramatises ideas." (Eisner, 2008b). The key to designing comics lies in how the spatial arrangement of panels (separated by white space, called the gutter) allows the reader to deduce a temporal or causal sequence (Foulsham et al., 2016; Mandler and Johnson, 1977). The sequence forms a visual narrative that needs 'reading' (Mandler and Johnson, 1977). This 'reading' demands readers to interpret how images represent events (Johnson-Laird et al., 1992) to infer causal sequences (Wilensky, 1983). From this perspective, we use text analysis technology to identify content that can undergo visualization in comics. Of particular importance is the fact that we have developed systems capable of summarizing text into understandable visual narratives. As a result, information can be conveyed with minimal language knowledge, or key concepts in the content can be rapidly assimilated.

Regarding text analysis, the first step is extracting the text from its original source in preparation for the analysis. In this project, we assume the original source to be in the form of webpages. We can apply a Python-based web crawler technology to extract text from webpages (Kausar et al., 2013). For the web crawler extraction, we demonstrate the extraction principle in the following figure:

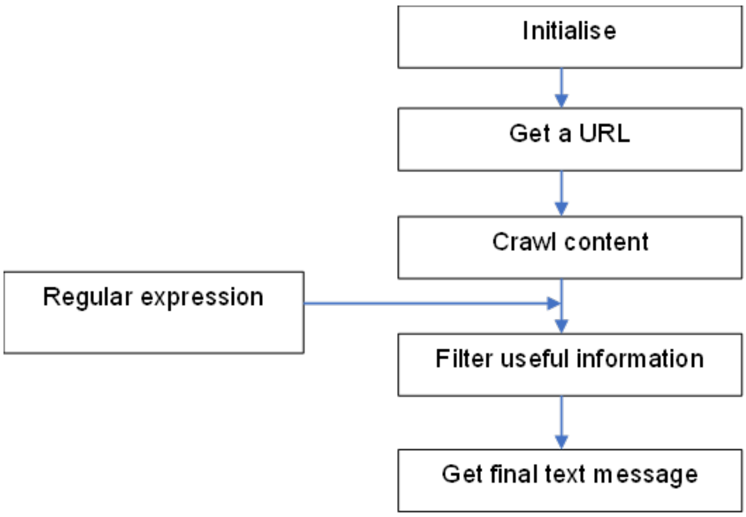


Figure 16. The working process of the crawler system

In the crawler system, we first need to determine what information to extract from the web page. This operation requires us to first inspect the design rules in the webpage's background to design the necessary regular expressions. In a typical web page, text is embedded between the three patterns <p></p>, , and <h3></h3>. Therefore, we need to use these three patterns to design regular expressions to extract the body text information that we need.

The text summarization method we discuss in this chapter is based on natural language processing. We used a natural language toolkit to study the features of the text's terms and their dependency relationship (Loper and Bird, 2002; Tas and Kiyani, 2007). The part of speech of a vocabulary term is crucial to natural language analysis. It not only determines the function of a single word in a phrase but also lays the foundation for the interdependence of words. Using Python for natural language analysis and processing on text data boosts and simplifies the performance of interactive systems (Van Rossum and Drake, 2003; Loper, 2004; VanRossum and Drake, 2010).

3.1 Requirements and Specification for Comic System

As previously said, comics convey narrative texts through images. Consequently, the elements found in comics also exist in the story framework. The basic components of comic design are subject, action, object, and result. These characteristics are also crucial to the operation of a system for generating comics. Once we store all the event-related vocabulary in the dictionary, the comic system can summarize the event system and generate comics based on the word's part of speech and the dictionary's key. The diagram below illustrates the operation of the comic system:

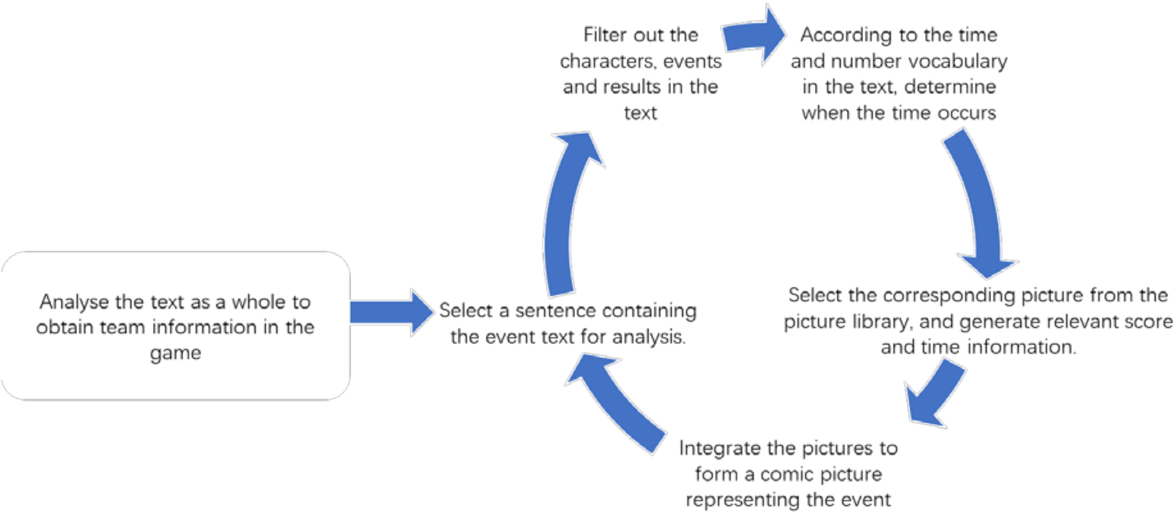


Figure 17. The operating process of the comic system

Our comic generation process involves the following steps (figure 18):

- Step1: Crawl the original report from web
- Step2: Text Analysis
- Step2: Create a dictionary based on story grammar
- Step3: Create an image library corresponding to the event
- Step4: Generate image corresponding to the event and combine it into a comic

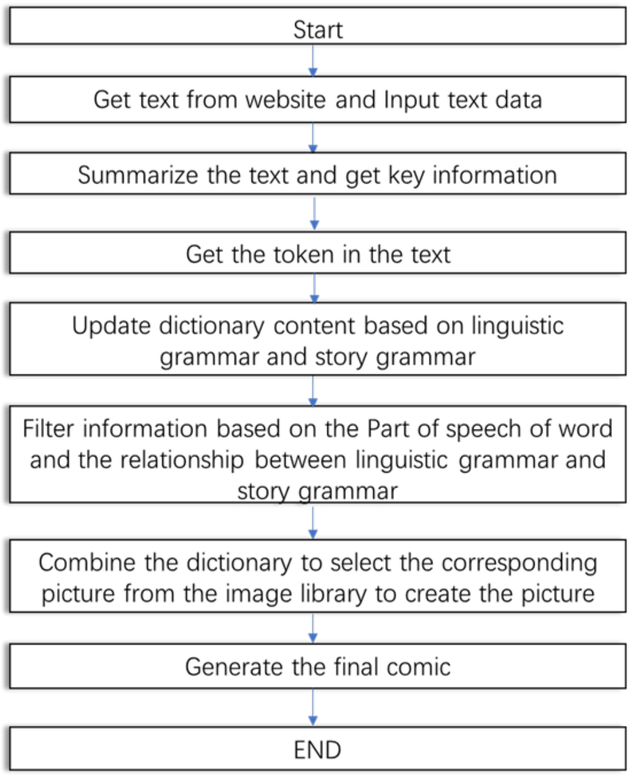


Figure 18. comic generation process

Almost all game reports contain a significant amount of redundant information. These reports aid the reader in extracting additional information from the text. For this project's system, however, the required text in the comic creation method must be a text containing all the pertinent information about the current game. In addition, the language must be brief enough to summarise all the relevant information regarding the current game in fewer words. As a result, for the original report, we devised an acceptable programme to summarise the original language and extract the text that the system could directly use.

The analysis and processing of the text are based on the grammar of the story. The grammatical structure is shown in the figure below:

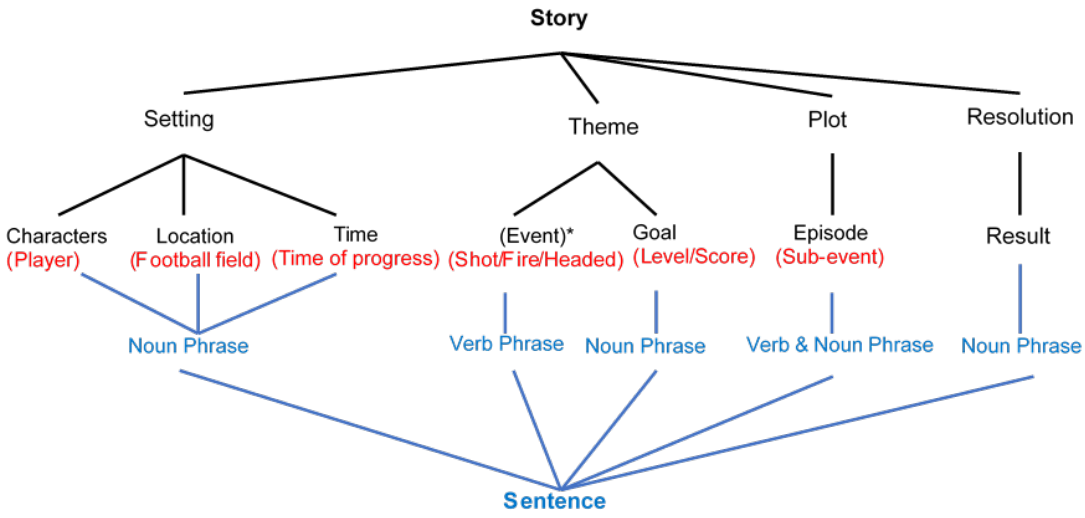


Figure 19. The story grammar of a soccer match

Step 1 Extract Text from Webpage

The first step in our system is to crawl the original information of the report of a soccer game from the internet. Here, we need to obtain the URL link to the webpage and then observe the background code of the web page to obtain the expression of the text embedding code to design the corresponding regular expression to extract the text. We cannot create a system to extract original text information unless we have obtained the original text data. The following is the code used to crawl the original report's data:

```
def get_html(url):
    headers = {
        'User-Agent':'Mozilla/5.0(Macintosh; Intel Mac OS X 10_11_4)\
        AppleWebKit/537.36(KHTML, like Gecko) Chrome/52 .0.2743. 116 Safari/537.36'

    }
    response = requests.get(url,headers = headers)
    html = response.text
    return html
reg = r'<p>(.*)</p>|<li>(.*)</li>|<h3>(.*)</h3>'
reg_ques = re.compile(reg)      #编译一下正则表达式，运行的更快
queslist = reg_ques.findall(get_html('https://www.skysports.com/football/everton-vs-c-palace/408235'))
```

Figure 20. the code for crawl the report information from the website.

Figure 21 demonstrates that the text information collected in this manner contains a significant amount of unneeded script information.

```
'#{player}', ''), ('', '#{player}', ''), ('', '<a href="/get-sky"  target="_blank" >Get Sky
Sports</a></li><li><a href="https://www.nowtv.com/promo/sky-sports?
dcmp&#x3D;ilc_SSNTV_skysports_hardcode_sportspass"  target="_blank" >Get a Sky Sports
Pass</a>', ''), ("Everton maintained their recent good form as they overcame struggling
Crystal Palace 3-1 at Goodison Park on Saturday lunchtime to move up to seventh in the
Premier League.</p><p>Bernard opened the scoring after just 18 minutes with a well-struck
volley, the Brazilian's third goal of the season on his 50th league appearance for
Everton.", '', ''), ('', '<strong><a href="https://www.skysports.com/football/everton-vs-c-
palace/teams/408235">How the teams lined up</a></strong></li><li><strong><a href="https://
www.spreaker.com/user/skysports/pl-weekend-preview-07-feb">Listen: Weekend Preview
podcast</a></strong></li><li><strong><a href="https://www.skysports.com/premier-league-
table">Premier League Table</a> | <a href="https://www.skysports.com/premier-league-
fixtures">Fixtures</a> | <a href="https://www.skysports.com/football/news/11661/11756515/
premier-league-top-goalscorers-201920">Top scorers</a></strong></li><li><strong><a
href="https://www.skysports.com/football/news/11661/11926353/premier-league-winter-break-
how-is-your-club-spending-it">PL winter break: Who\'s going where?</a></strong>', ''),
("However, despite dominating the first half, Palace - who fielded their oldest-ever
```

Figure 21. The confusing text information (because it contains web scripts)

Therefore, we must configure regular expressions in the crawler system to eliminate superfluous script data.

```
s =re.sub('<a .*?>(.*?)</a>','',queslist)
s =re.sub('</li><li>','. ',s)
s =re.sub('#{player}','',s)
s =re.sub('<strong></strong>','',s)
s =re.sub('<strong>','',s)
s =re.sub('</strong>','',s)
s =re.sub('</h3>','',s)
s =re.sub('<ul>','',s)
s =re.sub('<li>','',s)
#s =re.sub(' |  | ','',s)
s =re.sub('</p><p>','',s)
s =re.sub('</li>','',s)
s =re.sub('</ul>','',s)
s =re.sub('<h3>','',s)
s =re.sub('<p>','',s)
s =re.sub('</p>','',s)
s =re.sub('</em>','',s)
s =re.sub('<em>','',s)
```

Figure 22. Regular expression code

Following the application of regular expressions, the original text is produced:

Everton maintained their recent good form as they overcame struggling Crystal Palace 3-1 at Goodison Park on Saturday lunchtime to move up to seventh in the Premier League. Bernard opened the scoring after just 18 minutes with a well-struck volley, the Brazilian's third goal of the season on his 50th league appearance for Everton. However, despite dominating the first half, Palace - who fielded their oldest-ever starting line up in Premier League history - drew level six minutes after half-time thanks to Christian Benteke's first top-flight goal since last April, only for Richarlison to fire the hosts back in front just before the hour-mark. Dominic Calvert-Lewin's late header confirmed Everton's victory, the striker's sixth goal in eight appearances under new boss Carlo Ancelotti. As a result, Everton move up to seventh in the table, while Roy Hodgson's team stay 14th, but after just one win in their last 11 games, the Eagles are just six points from safety having now played a match more than their relegation rivals. 'How Everton jumped up to seventh'. Everton came into the game on the back of just one defeat in their previous 10 matches, form that had seen Ancelotti's side surge up the table, and that was borne out in their early dominance. Lucas Digne tested Vicente Guaita with a third-minute free kick and while Patrick van Aanholt did see his deflected drive hit the outside of the post, the home side soon found themselves ahead. Theo Walcott showed quick feet to skip past Van Aanholt's attentions on the right, before whipping in an inviting ball that the unmarked Bernard swept past Guaita with a cushioned timed volley. Against the run of play, however, the visitors levelled matters when Benteke collected Wilfried Zaha's through-ball and somehow fired a shot through Jordan Pickford to end a barren 18-match scoreless run in the league. It is not a goal the England goalkeeper will look back on with any fondness though. Not that the scores stayed level for long, with the hosts retaking the lead seven minutes later after a piece of individual brilliance from Richarlison, the Brazil forward's 11th goal of the campaign. And despite Benteke hitting a post with a close-range header, Everton sealed the win with Calvert-Lewin's 88th-minute header, the 22-year-old's 13th goal of the season maintaining his side's hopes of European football next season with the Toffees now sitting just a point off Tottenham in fifth. 'Opta stats'. Only Liverpool (24) have won more Premier League points than Everton (17) since Carlo Ancelotti's first game in charge of the Toffees in December. Crystal Palace are without a clean sheet in 10 Premier League games, the longest current run in the competition. However, this was just the second time the Eagles have conceded 3+ goals in a Premier League match this season (also 0-4 vs Spurs in September 2019). Crystal Palace named their oldest ever starting XI in the Premier League (30y 101d). None of the last 18 occasions a team has named a side with an average age of 30+ years in the competition has won (D3 L15), with Man City against Southampton in April 2017 the last such victory. Everton's Theo Walcott has as many goal involvements in his last two Premier League games (1 goal, 1 assist) as he did in his previous 20 in the competition (also 1 goal, 1 assist). Managers Carlo Ancelotti: "I prefer to win than to have a good performance and lose. It is normal. The result is really important for us after a performance that was not so good for 60 minutes. "In this moment we believe in what we are doing and that is the most important part. "The work everyone has done was really good: players, club, staff, everyone and now our table is good and we can dream to fight for the Europa League in the next games. "Roy Hodgson: "I thought we played well for large parts of the game. Not right at the start but neither side were they playing well," he said. "For 20-25 minutes there was not much going on but after that I thought we started to play much better. "We started the second half extremely well and when we got the equaliser I thought it was a fair reflection. "It is a defeat that hurts us badly. " Man of the Match - Richarlison. The Brazil international was the subject of reported interest from Barcelona in his services in January, and on the basis of this energetic display, you can see why. The forward was a constant thorn in the Palace defence all afternoon long, perhaps best illustrated by the 22-year-old's crucial strike that gave his side the lead again minutes after the visitors had equalised. Latching on to Calvert-Lewin's clever flick-on, Richarlison still had a lot of work to do, only for the striker to turn the experienced Gary Cahill inside out, before his curling effort found the bottom righthand corner of the net. "What's next?", Everton face Arsenal at the Emirates - a match you can see live on Sky Sports Premier League - although not till Sunday February 23 after their midseason break. Meanwhile, Palace take on Newcastle at Selhurst Park, with the Eagles having to wait until Saturday February 22 for their next league outing.

Figure 23. The original text after regular expression code

Further refinement of the code, allows the text summary system to obtain the most simplified summary text.

Step 2 Natural language processing and analysis

We input the original text is into the Python Spacy natural language processing library. This decomposes the text into Parts of Speech (see chapter 1). After analysing the sentences in the original text, the results obtained are shown in the following figure:

Bernard	0	bernard	False	False	Xxxxx	PROPN	NNP
opened	8	open	False	False	xxxx	VERB	VBD
the	15	the	False	False	xxx	DET	DT
scoring	19	scoring	False	False	xxxx	NOUN	NN
after	27	after	False	False	xxxx	ADP	IN
just	33	just	False	False	xxxx	ADV	RB
18	38	18	False	False	dd	NUM	CD
minutes	41	minute	False	False	xxxx	NOUN	NNS
with	49	with	False	False	xxxx	ADP	IN
a	54	a	False	False	x	DET	DT
well	56	well	False	False	xxxx	ADV	RB
-	60	-	True	False	-	PUNCT	HYPH
struck	61	strike	False	False	xxxx	VERB	VBN
volley	68	volley	False	False	xxxx	ADJ	JJ
,	74	,	True	False	,	PUNCT	,
the	76	the	False	False	xxx	DET	DT
Brazilian	80	brazilian	False	False	Xxxxx		
's	89	's	False	False	'x	PART	POS
third	92	third	False	False	xxxx	ADJ	JJ
goal	98	goal	False	False	xxxx	NOUN	NN
of	103	of	False	False	xx	ADP	IN
the	106	the	False	False	xxx	DET	DT
season	110	season	False	False	xxxx	NOUN	NN
on	117	on	False	False	xx	ADP	IN
his	120	-PRON-	False	False	xxx	ADJ	PRP$
50th	124	50th	False	False	ddxx	ADJ	JJ
league	129	league	False	False	xxxx	NOUN	NN
appearance	136	appearance	False	False	xxxx		
for	147	for	False	False	xxx	ADP	IN
Everton	151	everton	False	False	Xxxxx	PROPN	NNP
.	158	.	True	False	.	PUNCT	.

Figure 24. The results of language features

The features include (from left to right) text, index value (found in the original text), lemma, punctuation, spacing, shape, part of speech, and tag.

After identifying the Part of Speech for each word in the original text, the vocabulary was screened based on associated linguistic grammar and story grammar. Following are the function code and results:

```
f = open('C:/Users/DELL/Desktop/dic.txt', 'a')
for i in range(leng):
    print(list(docCR.sents)[i])
    sentence=list(docCR.sents)[i].text
    for token in list(docCR.sents)[i]:
        #print(token.pos_)
        if token.pos_ == 'PROPN':
            print(token.text)
            propnlist.append(token.text)
            propn.append(token.text)
        if token.pos_ == 'VERB':
            print(token.text)
            verblist.append(token.text)
            verb.append(token.text)
        if token.pos_ == 'NOUN':
            print(token.text)
            nounlist.append(token.text)
            noun.append(token.text)
    resultlist.append(str(sentence)+"    "+'PROPN:  '+str(propnlist)+"    "+'VERB:  '+str(verblist)+"    "+'NOUN:  '+str(nounlist))
    nounlist=[]
    verblist=[]
    propnlist=[]
    f.write(str(resultlist))
    del resultlist[0]

    f.write('\n')
f.write('PROPN:    '+str(set(propn)))
f.write('\n')
f.write('NOUN:    '+str(set(noun)))
f.write('\n')
f.write('VERB:    '+str(set(verb)))
f.write('\n')
f.close()
```

Figure 25. The function code to screen out the required words

```
PROPN:
{'April', 'Richarlison', 'League', 'Lewin', 'Brazilian', 'Brazil', 'Christian', 'Carlo', 'Tottenham', 'Everton', 'Bernard', 'Palace', 'Calvert', 'Toffees', 'Premier', 'Ancelotti', 'Benteke'}
NOUN:
{'who', 'hopes', 'scoring', 'time', 'line', 'brilliance', 'front', 'league', 'half', 'goal', 'point', 'minutes', 'striker', 'boss', 'campaign', 'header', 'level', 'season', 'post', 'hour', 'minute', 'appearances', 'flight', 'win', 'range', 'thanks', 'history', 'football', 'forward', 'side', 'mark', 'hosts', 'appearance', 'scores', 'victory', 'piece'}
VERB:
{'starting', 'drew', 'struck', 'sitting', 'retaking', 'maintaining', 'fire', 'fielded', 'confirmed', 'opened', 'hitting', 'sealed', 'stayed', 'dominating'}
NUM:
{'six', '18', 'eight', '22-year', 'seven'}
```

Figure 26. The result of the required words

Moreover, the interdependence of words is an essential component parameter for our system's architecture. Through the code depicted in Figure 27, the process can utilise the direction, similar to vector data, to process and analyse text data. Consider the first clause as an example. Figure 28 displays the resulting visual relationship diagram.

```
for token in doc:
    print("{0}({1})      <--{2}--<          {3}({4})".format(token.text, token.tag_, token.dep_, token.head.text, token.head.tag_))
displacy.render(doc, style='dep', jupyter=True, options={'distance': 120})
```

Figure 27. The functional statements about the dependencies

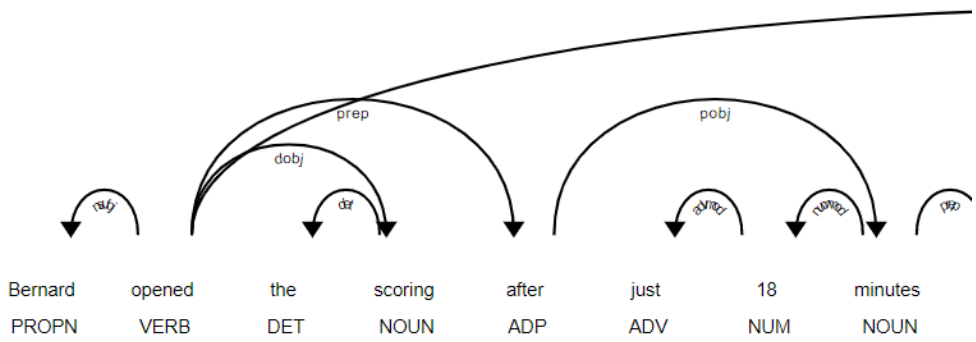


Figure 28. The example about visualisation of the dependencies

Corresponding to the preceding clause and the narrative constructions outlined in the introduction(figure 11). The association is depicted in the figure below:

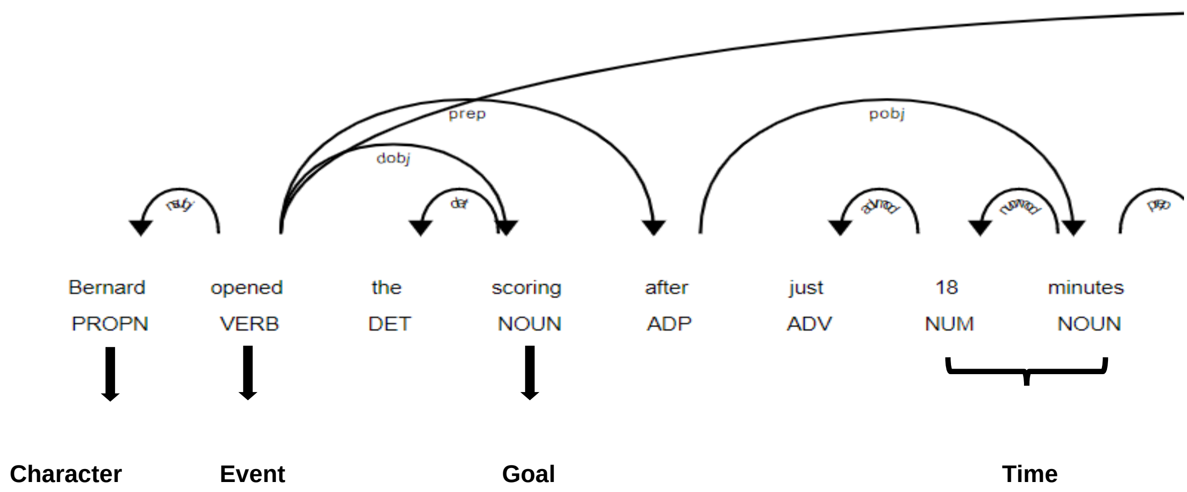


Figure 29. The correlation diagram in practical application

As a result of the foregoing natural language analysis procedure, the system has collected a variety of text information parameters, and a dictionary system has been built.

Step 3 The Creation of Dictionary

Once we retrieve the essential components of the story, the critical next step is to develop the lexicon. Because the system is based on story grammar, we must design the dictionary to fit within the story grammar framework.

During the examination of a plot, different words often signify the same action. Therefore, determining whether distinct words correspond to the same visual element is a vital design step. In the initial design of the dictionary, we used WordNet, a large lexical database, to create the dictionary (Miller, 1995). This method uses the WordNet thesaurus to traverse the synonym set related to the specified vocabulary and print the output result. Thus, we input the extracted keywords (such as nouns and verbs) into the WordNet thesaurus one by one using a for loop in Python to make calls to WordNet, obtaining the synonyms related to it. The following figure shows the dictionary constructed in this way.


```
{
    "NOUN": {
        "say": [
            "say",
            "enjoin",
            "sound_out",
            "allege",
            "tell",
            "read",
            "state",
            "aver",
            "suppose",
            "pronounce",
            "enounce",
            "articulate",
            "enunciate",
            "order"
        ],
        "angle": [
            "weight",
            "angle",
            "tip",
            "fish",
            "Angle",
            "lean",
            "slant",
            "tilt"
        ],
        "time": [
            "metre",
            "time",
            "prison_term",
            "clip",
            "sentence",
            "clock",
            "fourth_dimension"
```


Figure 30. The synonym dictionary

This synonym dictionary is just the most common word-level dictionary and doesn't concern the grammatical structure of the text. Furthermore, because of its multilayered nature, the system can't quickly read and recall it during operation. Therefore, we can't directly apply the synonym dictionary extracted from WordNet to our system. We must design the dictionary to conform to the grammatical structure of the story. As previously shown in Figure 11, we create dictionaries by linking linguistic grammars and story grammars. First, we need to create the dictionary keys according to the story grammar structure, which will facilitate the classification of the extracted feature words. The keys to a dictionary based on story grammar are set to event_num, event_person, event_goal, event_team and event_result. Then we can use Named Entity Recognition to identify the names of players and teams, so they can be classified into the key event_person directory. Here, we usually mark the person's name with "PERSON," and we mark the team with "ORG". We need to manually arrange the team after the person's name so that we can mark the person's team in the event with a secondary value. At the same time, we classify the team name into event_team. Then, the system analyzes the part of speech of the vocabulary in the sentence. For the description of the result, the sentence usually uses nouns. Therefore, we classify words whose part of speech is a noun into event_result, such as level, score, etc. Finally, verbs usually express actions related to events, so for verb vocabulary, we attribute it to event_goal, such as heading or kicking the ball, etc. Words related to numbers have the "NUM" tag, so we classify them under the event_num key and convert words and Arabic numerals to them. In addition, the dictionary must have time-related keys to identify the story's chronology.

For the establishment of dictionaries, we can create them through the spacy library. First we can manually create an empty dictionary containing all keys. Then use the json library to load the dictionary, which will be able to continuously add the required values through the .append() function in the later stage.

```
with open(json_file_path, 'r') as file:
    data = json. load(file)
```

Subsequently, we use en_core_web_sm in the spacy library to analyze the characteristics of each word in the text, and add the words that meet the characteristics to the corresponding key. For example, we can add verbs representing events to the key 'event_goal' by the following code.

```
for token in doc.ents:
    token. pos_ == 'VERB':
        data['event_goal'].append(ent.text)
```

In the same operation, we only need to change the value of the key in data[key] to add the corresponding vocabulary to the corresponding key. Therefore, after traversing the full text, we can get the corresponding dictionary. Here, there is a step that we need to modify manually, that is to add the team we belong to after the value in the key 'event_person', for example: "SergioAguero":"Manchester City". The resulting dictionary is shown below.:


```
{   "half-":"0.5",
    "end":"1",
    "time": "90",
    "event_num":
    {
        "Five":"5",
        "15":"15",
        "64":"64",
        "36":"36",
        "59":"59",
        "seven":"7",
        "six":"6",
        "18":"18",
        "eight":"8",
        "88th":"88",
        "fifth":"5",
        "59th":"59",
        "four":"4",
        "12th":"12",
        "69th":"69",
        "first":"1",
        "52nd":"52",
        "58th":"58",
        "half":"0.5"
    },
"event_person":
    {
        "SergioAguero":"Manchester City",
        "RoDertoFirmino":"Liverpool",
        "Sane":"Manchester City",
        "Bob":"Manchester City",
        "RoDerto":"Manchester City",
        "Sandy":"Liverpool",
        "Ben":"Liverpool",
        "Sagir":"Manchester City",
        "Pickford":"Everton",
        "Coleman":"Everton",
        "Keane":"Everton",
        "Mina":"Everton",
        "Digne":"Everton",
        "Schneiderlin":"Everton",
        "Walcott":"Everton",
        "Sigurdsson":"Everton",
        "Bernard":"Everton",
        "Richarlison":"Everton",
        "Lewin":"Everton",
        "Guaita":"Crystal Palace",
        "Ward":"Crystal Palace",
        "Tomkins":"Crystal Palace",
        "Cahill":"Crystal Palace",
        "VanAanholt":"Crystal Palace",
        "Ayew":"Crystal Palace",
        "McCarthy":"Crystal Palace",
        "Milivojevic":"Crystal Palace",
        "McArthur":"Crystal Palace",
        "Zaha":"Crystal Palace",
        "Benteke":"Crystal Palace",
        "Wilshere":"Bournemouth",
        "King":"Bournemouth",
        "Nichols":"Crawley Town",
        "Eisa":"Scunthorpe United",
        "Green":"Scunthorpe United",
        "O'Donnell":"Bradford City",
        "Shelvey":"Newcastle United",
        "Riedewald":"Crystal Palace",
        "Cahill":"Crystal Palace",
        "Macrae":"Brora",
        "Maclean":"Brora",
        "Berra":"Hearts",
        "Yates":"Blackpool",
        "Garbutt":"Blackpool",
        "Ward":"Peterborough United"
    },
"event_goal":
    [
        "lashed",
        "sent",
        "firing",
        "headed",
        "crashed",
        "put",
        "fired",
        "scored",
        "get",
        "socre",
        "steal",
        "tried",
        "injured",
        "fell",
        "held",
        "tackled",
        "ruled",
        "confirmed",
        "hitting",
        "sealed",
        "opened",
        "retaking",
        "confirmed",
        "sealed",
        "maintaining",
        "sitting",
        "dominating",
        "stayed",
        "starting",
        "opened",
        "fire",
        "hitting",
        "struck",
        "fielded",
        "retaking",
        "sealed",
        "struck",
        "equaliser",
        "slot",
        "lash",
        "restored",
        "flick",
        "took",
        "volleyed",
        "saw",
        "wrapped",
        "level"
    ],
"event_team":
    [
        "Liverpool",
        "Manchester City",
        "Crystal Palace",
        "Everton",
        "Crawley Town",
        "Bournemouth",
        "Scunthorpe United",
        "Bradford City",
        "Newcastle United",
        "Hearts",
        "Brora",
        "Blackpool",
        "Peterborough United"
    ],
    "event_result":
    [
        "level",
        "shot",
        "front",
        "say",
        "delight",
        "point",
        "score",
        "minutes",
        "leg",
        "ground",
        "football",
        "offside",
        "net",
        "ball",
        "closing",
        "seconds",
        "scoring",
        "forward",
        "mark",
        "post",
        "scores",
        "win",
        "flight",
        "goal",
        "lead",
        "victory",
        "home",
        "parity",
        "turnaround",
        "lead",
        "opener",
        "hosts",
        "equaliser",
        "start"
    ]
}
```


Figure 31. The dictionary based on story grammar

As the user cannot access the system's history, upgrading the dictionary becomes a critical issue. We designed a user interface for the dictionary, making it easier for users to update the dictionary (see Figure 32). Like the previous text analysis work, we can add a large number of similar words to the dictionary section. To allow the system to recognize nouns, name prepositions, and verbs, I included functional phrases in the software:

if token.pos_=="VERB":

if token.pos_=="NOUN":

if token.pos_=="PROPN":

After the vocabulary update mechanism completes the recognition, the user can choose whether to contribute to the new dictionary. Once the system uploads all selected codes, the dictionary update system can fully cover all essential textual information.

For example, the text is: "***Bernard opened the scoring after 18 minutes.***". The user will access the related interactive page after launching the dictionary system (as shown in figure 32).

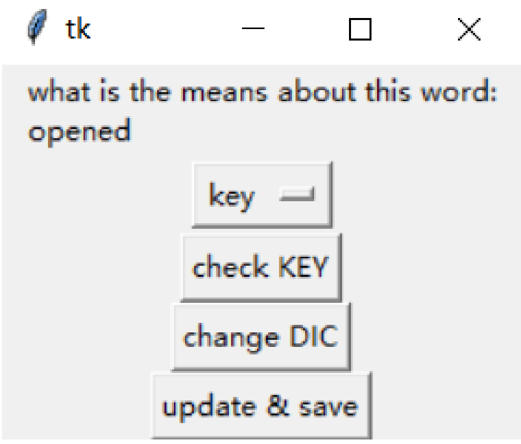


Figure 32. User Interface of Dictionary

Figure 33. The workflow of the dictionary update system

On an interactive page, the user can view the vocabulary extracted by the present system, and select the key that must be added, and save it. By interacting with the dictionary's user interface, as depicted in figure 34, the user can add words to the dictionary.

```
    "event_goal":[
        "lashed",
        "sent",
        "firing",
        "headed",
        "crashed",
        "put",
        "fired",
        "scored",
        "get",
        "socre",
        "steal",
        "tried",
        "injured",
        "fell",
        "held",
        "tackled",
        "ruled",
        "get"
    ],
```

```
    "event_goal":[
        "lashed",
        "sent",
        "firing",
        "headed",
        "crashed",
        "put",
        "fired",
        "scored",
        "get",
        "socre",
        "steal",
        "tried",
        "injured",
        "fell",
        "held",
        "tackled",
        "ruled",
        "get",
        "opened"
```

Figure 34. Comparison of the content of the old and new dictionary

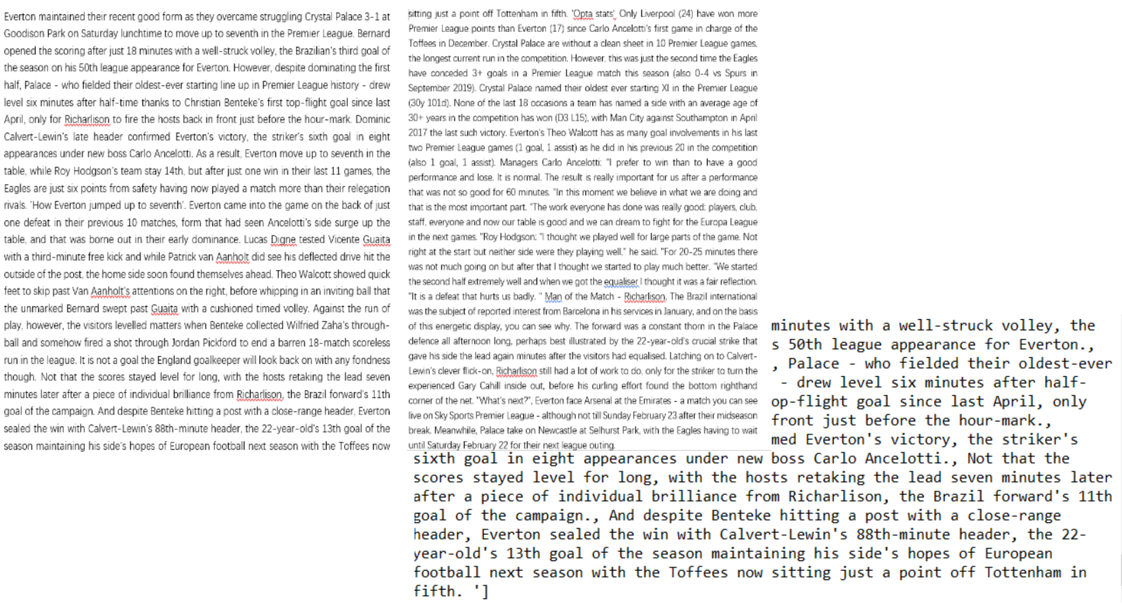
By comparing the dictionary's contents before and after the system is operating, we can determine that all of the crucial information from the original text that must be put to the dictionary has been added. Therefore, it appears that the existing operation of the dictionary update system satisfies the current design specifications.

Having obtained the relevant dictionaries, we can use the system to simplify the extracted original text to obtain the simplest text information.

```python
for sent in doc.sents:
    for token in sent:
        if token.pos_=="VERB":
            verbtext=token.text
#             for token in sent:
#                 if token.text=="minutes":
#                     if sent not in currentevent:
#                         currentevent.append(sent)
            if token.dep_ !="conj":
                for k, v in temp.items():
                    if token.text in v:
                        if k=="event_goal":
                            for token in sent:
                                if token.head.text==verbtext and token.dep_=="dobj":
                                    if sent not in currentevent:
                                        currentevent.append(sent)
                            if token.text in v:
                                if k=="event_result":
                                    if koen.dep_=="amod":
                                        if sent not in currentevent:
                                            currentevent.append(sent)
        if token.pos_=="ADJ":
            adjtext=token.text
#             for token in sent:
#                 if token.text=="minutes":
#                     if sent not in currentevent:
#                         currentevent.append(sent)
            for k, v in temp.items():
                    if token.head.text in v:
#                         print(1231213212)
                        if k=="event_goal":
#                             for token in sent:
                                if token.dep_=="dobj":
                                    if sent not in currentevent:
                                        currentevent.append(sent)

        if token.pos_=="NOUN":
            tokentext1=token.head.text

            if token.dep_=="nsubj":
                for k, v in temp.items():
                    if tokentext1 in v:

                        if sent not in currentevent:
                                currentevent.append(sent)
```

Figure 35. The summary system code

The improved code reveals that the mechanism for optimising the simplest texts is based on the part of speech, dictionary, and vocabulary reliance. Additionally, this is a crucial parameter required by the comic creation system in the future.


Everton maintained their recent good form as they overcame struggling Crystal Palace 3-1 at Goodison Park on Saturday lunchtime to move up to seventh in the Premier League. Bernard opened the scoring after just 18 minutes with a well-struck volley, the Brazilian's third goal of the season on his 50th league appearance for Everton. However, despite dominating the first half, Palace - who fielded their oldest-ever starting line up in Premier League history - drew level six minutes after half-time thanks to Christian Benteke's first top-flight goal since last April, only for Richarlison to fire the hosts back in front just before the hour-mark. Dominic Calvert-Lewin's late header confirmed Everton's victory, the striker's sixth goal in eight appearances under new boss Carlo Ancelotti. As a result, Everton move up to seventh in the table, while Roy Hodgson's team stay 14th, but after just one win in their last 11 games, the Eagles are just six points from safety having now played a match more than their relegation rivals. 'How Everton jumped up to seventh'. Everton came into the game on the back of just one defeat in their previous 10 matches, form that had seen Ancelotti's side surge up the table, and that was borne out in their early dominance. Lucas Digne tested Vicente Guaita with a third-minute free kick and while Patrick van Aanholt did see his deflected drive hit the outside of the post, the home side soon found themselves ahead. Theo Walcott showed quick feet to skip past Van Aanholt's attentions on the right, before whipping in an inviting ball that the unmarked Bernard swept past Guaita with a cushioned timed volley. Against the run of play, however, the visitors levelled matters when Benteke collected Wilfried Zaha's through-ball and somehow fired a shot through Jordan Pickford to end a barren 18-match scoreless run in the league. It is not a goal the England goalkeeper will look back on with any fondness though. Not that the scores stayed level for long, with the hosts retaking the lead seven minutes later after a piece of individual brilliance from Richarlison, the Brazil forward's 11th goal of the campaign. And despite Benteke hitting a post with a close-range header, Everton sealed the win with Calvert-Lewin's 88th-minute header, the 22-year-old's 13th goal of the season maintaining his side's hopes of European football next season with the Toffees now sitting just a point off Tottenham in fifth. 'Opta stats'. Only Liverpool (24) have won more Premier League points than Everton (17) since Carlo Ancelotti's first game in charge of the Toffees in December. Crystal Palace are without a clean sheet in 10 Premier League games, the longest current run in the competition. However, this was just the second time the Eagles have conceded 3+ goals in a Premier League match this season (also 0-4 vs Spurs in September 2019). Crystal Palace named their oldest ever starting XI in the Premier League (30y 101d). None of the last 18 occasions a team has named a side with an average age of 30+ years in the competition has won (D3 L15), with Man City against Southampton in April 2017 the last such victory. Everton's Theo Walcott has as many goal involvements in his last two Premier League games (1 goal, 1 assist) as he did in his previous 20 in the competition (also 1 goal, 1 assist). Managers Carlo Ancelotti: "I prefer to win than to have a good performance and lose. It is normal. The result is really important for us after a performance that was not so good for 60 minutes. "In this moment we believe in what we are doing and that is the most important part. "The work everyone has done was really good: players, club, staff, everyone and now our table is good and we can dream to fight for the Europa League in the next games. "Roy Hodgson: "I thought we played well for large parts of the game. Not right at the start but neither side were they playing well," he said. "For 20-25 minutes there was not much going on but after that I thought we started to play much better. "We started the second half extremely well and when we got the equaliser I thought it was a fair reflection. "It is a defeat that hurts us badly. " Man of the Match - Richarlison, The Brazil international was the subject of reported interest from Barcelona in his services in January, and on the basis of this energetic display, you can see why. The forward was a constant thorn in the Palace defence all afternoon long, perhaps best illustrated by the 22-year-old's crucial strike that gave his side the lead again minutes after the visitors had equalised. Latching on to Calvert-Lewin's clever flick-on, Richarlison still had a lot of work to do, only for the striker to turn the experienced Gary Cahill inside out, before his curling effort found the bottom righthand corner of the net. "What's next?", Everton face Arsenal at the Emirates - a match you can see live on Sky Sports Premier League - although not till Sunday February 23 after their midseason break. Meanwhile, Palace take on Newcastle at Selhurst Park, with the Eagles having to wait until Saturday February 22 for their next league outing.

```
minutes with a well-struck volley, the
s 50th league appearance for Everton.,
, Palace - who fielded their oldest-ever
 - drew level six minutes after half-
op-flight goal since last April, only
front just before the hour-mark.,
med Everton's victory, the striker's
sixth goal in eight appearances under new boss Carlo Ancelotti., Not that the
scores stayed level for long, with the hosts retaking the lead seven minutes later
after a piece of individual brilliance from Richarlison, the Brazil forward's 11th
goal of the campaign., And despite Benteke hitting a post with a close-range
header, Everton sealed the win with Calvert-Lewin's 88th-minute header, the 22-
year-old's 13th goal of the season maintaining his side's hopes of European
football next season with the Toffees now sitting just a point off Tottenham in
fifth. ']
```

Figure 36. The simplified summary after processing

Compared with the original text obtained in step 1, the processed text summary is shorter than original text and include all current events happened in match.

Step 4 The Creation of Image Library

In order to create a comic, the system needs to extract the corresponding images from the image library and place them in the correct position. During the creation of the comic, the system extracts multiple pictures from the library and places them in the correct position based on the results of the analysis and summary. With this step, the comic generation system can display the complete event through the images.

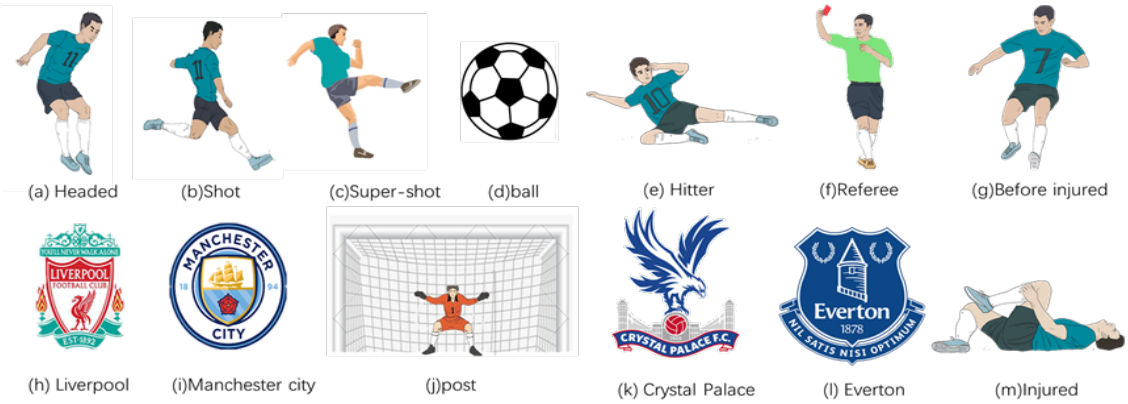


Figure37. The part of the pictures in images library for a football match

Step 5 Comic System

In this project, we needed to integrate comic generation with lexical dependencies and a dictionary based on story grammar to edit the image library and ensure the system accurately generated comics. We used Python statements to create the frame of the nine-square grid so that the system could place the called image in a specific position within the nine-square grid. This structure simplifies the system's task of building complete comics.

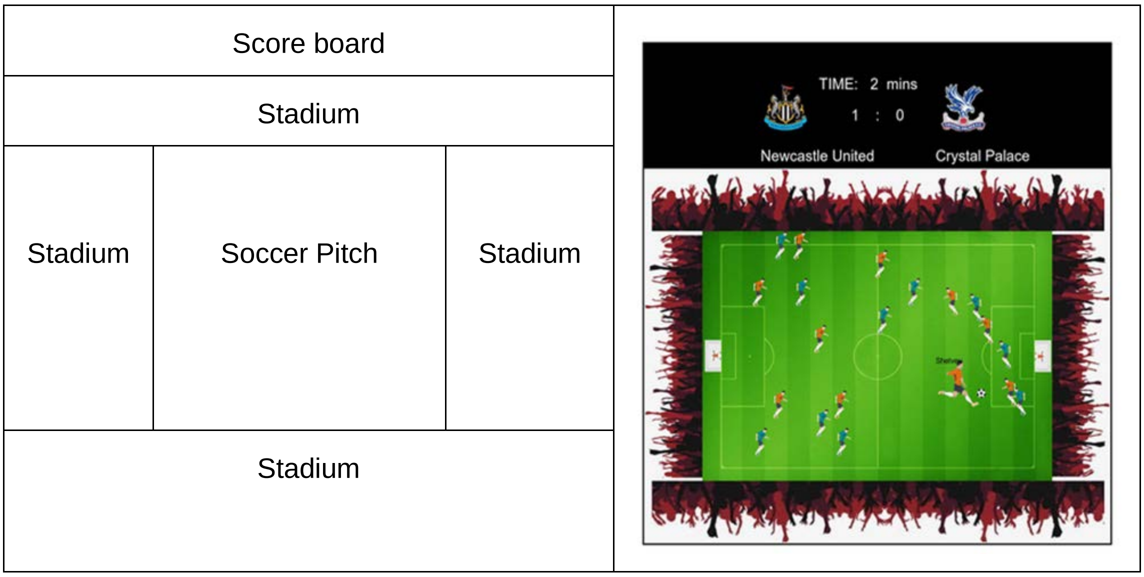


Figure 38. The layout of a panel (left) and example (right)

Step 5.1 Get the characteristic value in the report

After the text summarization system has processed the source text, each sentence in the summary text often reflects a whole event. Therefore, the machine will analyse each sentence separately.

The final feature item for the soccer game report is the names of the two teams. Therefore, the system must analyse the entire text in conjunction with the dictionary and assign various variables to the names of the competing teams. This operation is performed using the contestant's name and the value of the "event play" dictionary key.

```
for token in doc:
    if token.pos_=="PROPN":
        for k, v in temp.items():
                if token.text in v:
                    print(k)
                    if k =="event_person":
                        name=token.text
                        if team1=="a":
                           team1=temp[k][name]
                        if team1!="a" and team1!=temp[k][name]:
                           team2=temp[k][name]
```

Figure 39. The code for getting the characteristic value

The above code traverses all text information to allocate the names of the two teams in the game to separate variables. This allows the system to adapt to multiple reports without interference from other information in the dictionary. After assigning fixed feature values, subsequent operations use the part of speech of the word to determine the operative word in each phrase and the vocabulary dependence to identify the issuer and the result of the operation.

Step 5.2 Make event pictures

We selected a multiple-episode report from a newspaper related to a soccer game. The report reads:

> "*Five minutes before half-time, Sergio Aguero lashed home a magnificent near-post shot from a tight angle to put City in front. Roderto Firmino then headed Liverpool level after 64 minutes. But Sane crucially had the final say, firing home with a superb shot that crashed in off a post and sent the EtihadStadium wild with delight.* "

After the system filters out the verbs from the sentence, it identifies the position of the sentence in the dictionary. Once the verbs are recognized, they should be added to the dictionary, with the corresponding key value being "event goal". Additionally, the system must identify the adverb related to the verb. In the dependency connection of the sentence, "advmod" is the dependency tag corresponding to the adverbial, and the verb is the head text of the dependency; that is, the adverbial depends on the verb through the "advmod" dependency tag. In a soccer game, if the matching adverb to a verb is "home" or "post," it indicates a high likelihood of shooting action.

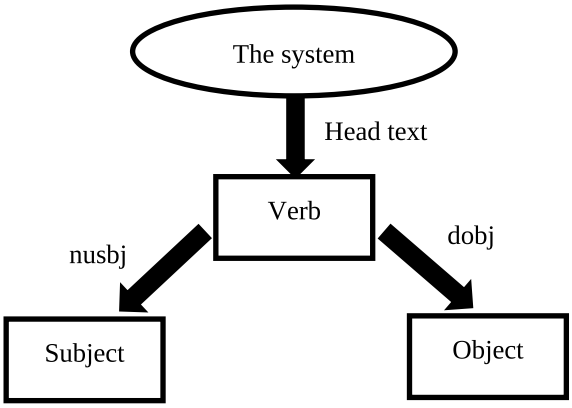


Figure 40: The event judgment process

The system then continues to detect and evaluate. When the system recognizes the presence of an adjective in a sentence, it determines whether the changed word, whose dependency is "amod", is "home" or "post". Once the judgement is made, the system makes a secondary judgement based on the directed dependency.

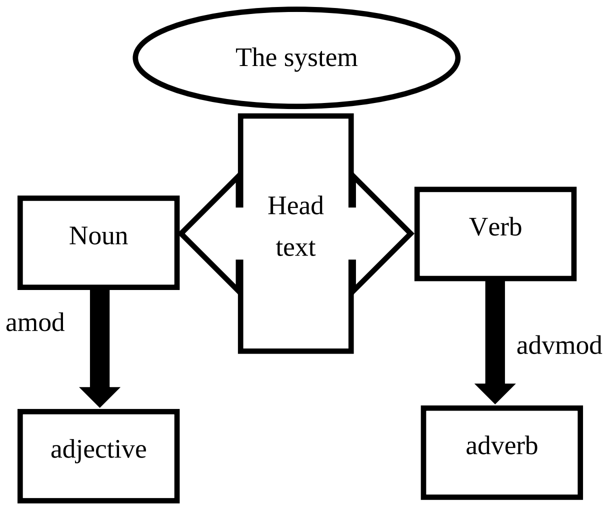


Figure 41: The more accurate recognition process

If the word being modified exists in a compound relationship and the vocabulary compounded with it is a noun indicating a shot, this indicates that the placement of the image representing the shooting action must be adjusted. If the detected adjective is "close" or similar, the image of the shooting action will be placed in the first row of the first column of nine squares. Otherwise, the system will place it somewhere on the third line. Since the sentence includes the event of a shot, the system places the image of the goal post in the first row and third column of the nine squares (the upper right square) and places the image of the soccer ball in the center of the nine squares. It seeks to match an image to an event.

From the text analysis discussed previously, we identify different events. Each of these events can be linked to an image in the system. For instance, if the system identifies a "VERB", it will be used to select an image related to the action. If the system detects a "PROPN", it will consult the dictionary to find out if it is a person's name and if this person belongs to a specific team. This will serve as the keyword for the image. The first sentence is an example:

> "*Five minutes before half-time, SergioAguero lashed home a magnificent near-post shot from a tight angle to put City in front*.".

The image created by the above steps is shown below:

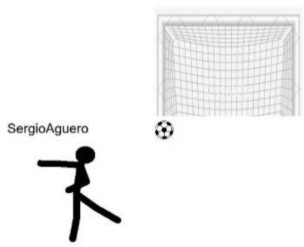


Figure 42. The picture result after first step

We must generate additional information related to an event once its construction is complete: In soccer games, the time and score are the most important supplementary information. Therefore, the system must append the time and score information for the current event to the previously generated image. Before this step can be processed, the system must establish a framework based on the team information in the dictionary.

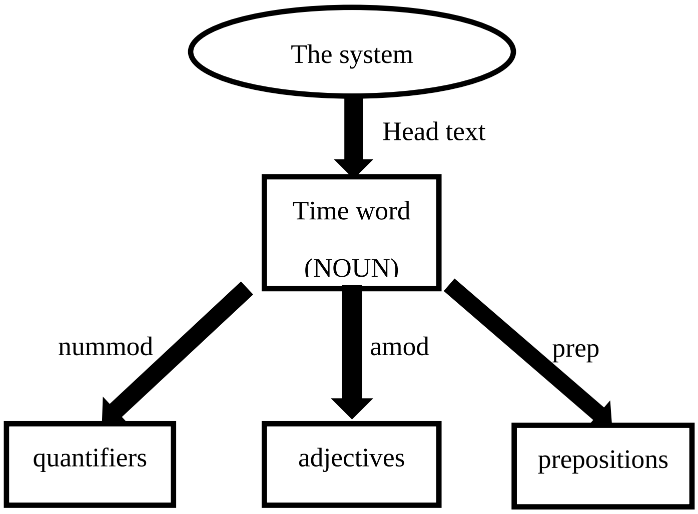


Figure 43: The Judgment of event time

This would produce the output:

TIME:
Score:
Liverpool 0
Manchester City 0

Figure 44. The infrastructure for additional information

It's important to note that the team name in "event_team" in the dictionary must match the team name in "event_person" exactly, including case. Otherwise, the system won't be able to link the person's name to the team.

After the basic framework has been built, the system processes and analyzes more textual input for the time element. The system first determines whether the text contains "time," "minute," and other time-related words. It then identifies time node A by finding the integer that modifies the time word. After identifying time node A, the algorithm looks for prepositions in the sentence, such as "before" and "after." When a preposition is found, the object of the preposition is identified as previously identified time-related vocabulary. Afterwards, if the detected vocabulary matches the corresponding number, the algorithm checks if there are more time-related vocabulary words before the preposition. Then, based on the preposition's properties, the system makes an exact computation by adding or subtracting to time node A again before completing the process. In the following frame, the time is displayed.

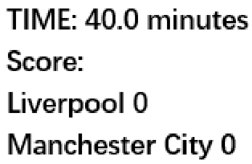
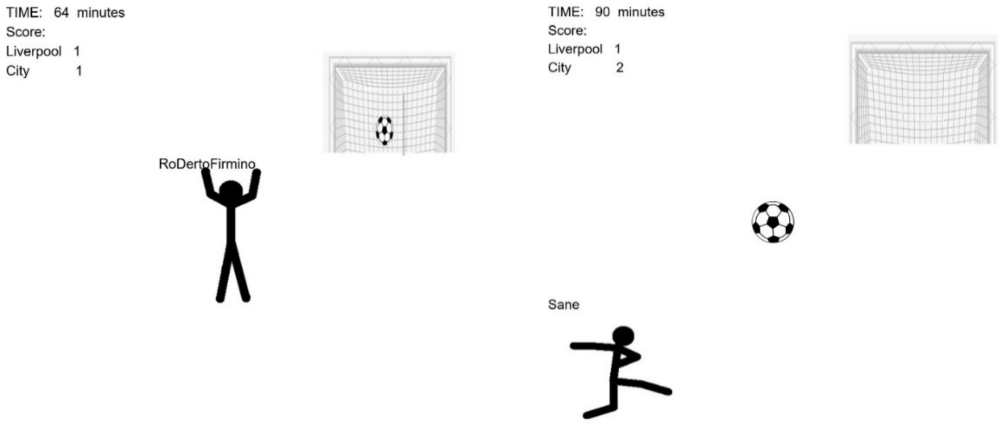


Figure 45. The time is shown in the frame

After processing the time information, the system evaluates the score information based on the nouns and verbs in the sentence. For instance, phrases like "front" and "level" carry the connotation of progression and when paired with the respective team name, the score of the related team can be inferred. Moreover, since human language often carries implicit meaning when there are no words in the sentence that express the meaning of progression, it's necessary to analyze the sentence components that can metaphorically express the result. For instance, the phrase "*sent the EtihadStadium wild with glee*" can also imply that a team scored a point from a bystander's perspective. In the initial sentence, the phrase "*put City in lead*" indicated that the team involving "SergioAguero" had scored a point. Finally, the result information implies that SergioAguero's team scored one point. The subsequent image outcomes are:

Figure 46. The image of first sentence

Using the same principle, the system processes the next two sentences and obtains the following two pictures.



(a) Second sentence (b) Third sentence

Figure 47. The image of another two sentence

After the images of the sentences are generated, the system sorts and merges them to create the final comic image. It's worth noting that each image's name corresponds to the appropriate period. Such naming criteria can assist the system in arranging all the images to ensure the story's accuracy.

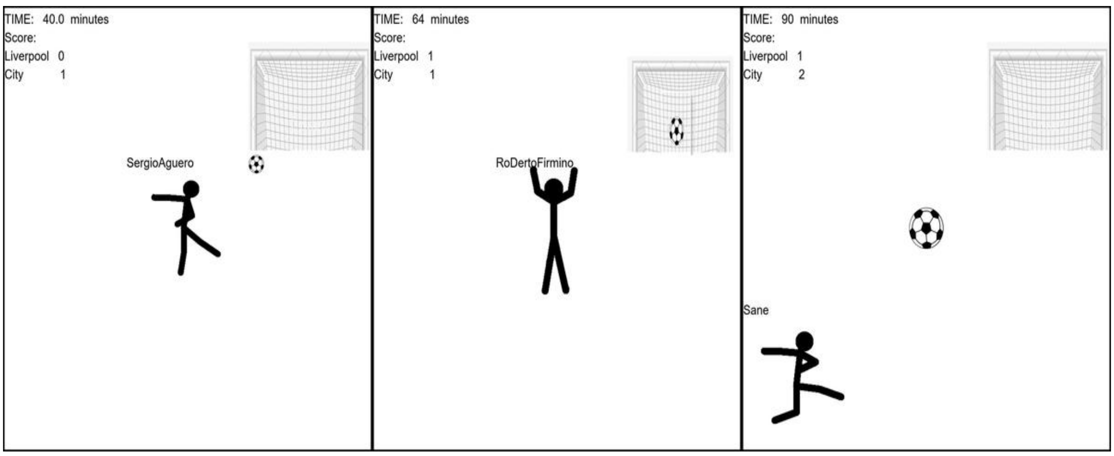


Figure 48. The finally result

The above layout is somehow not a general comic layout, so the layout of the generated image should be changed to the layout shown in figure 49.

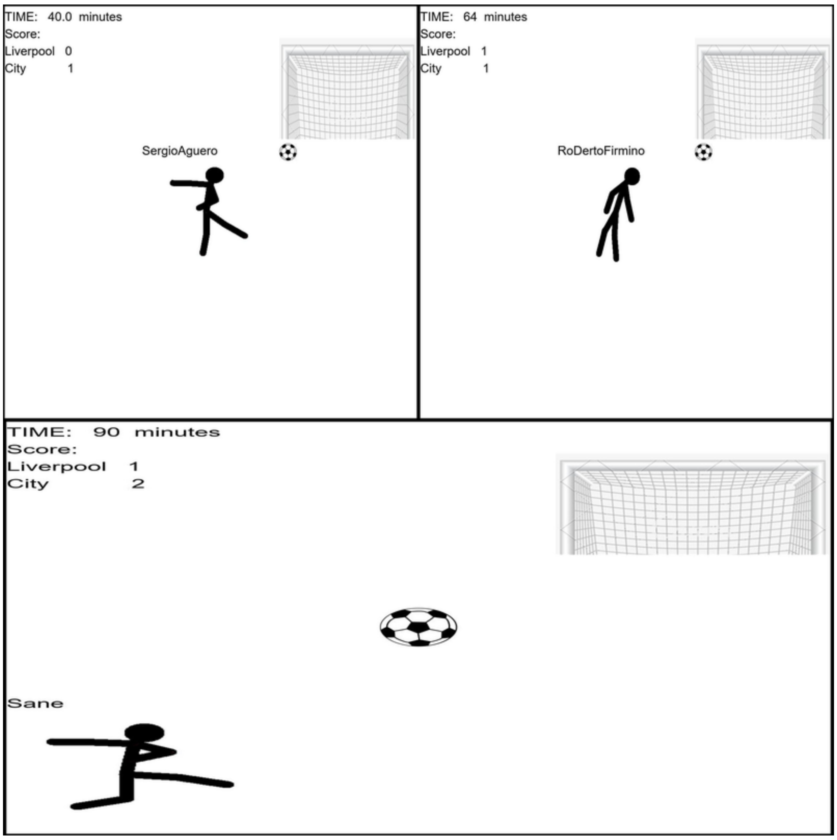


Figure 49. The finally result for general comic layout

The above comic is entirely composed of stick figures and its overall style is somewhat monotonous. Therefore, we need to adjust the arrangement to make the image more in line with the public's reading habits. Like the previous style, the base frame for each image in the comic is a nine-square grid frame. This type of frame makes it easier for the system to adjust the position of the image within the nine-square grids. However, we have created an additional information panel dedicated to displaying pertinent event details.

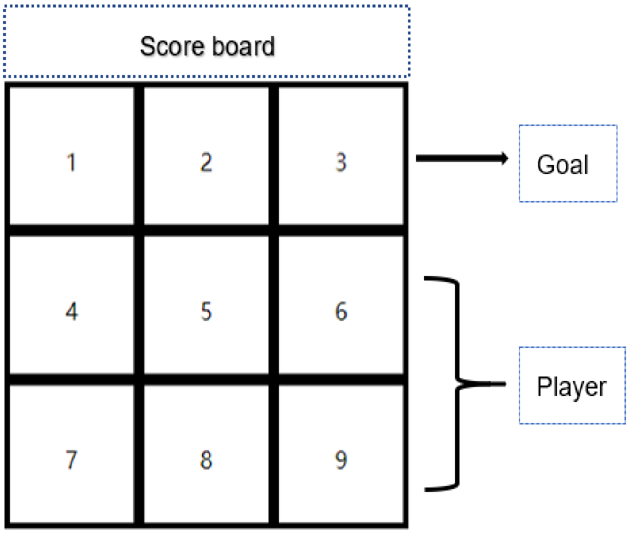


Figure 50. The layout of the nine-square grid

Step5.3 Make the comic

We use the summary text (Everton vs Crystal Palace) extracted in the previous section as material for a complete comic of a soccer game. We started by visualising the state-of-the-art dependencies of the first sentence (and also the first event) contained in the text information.

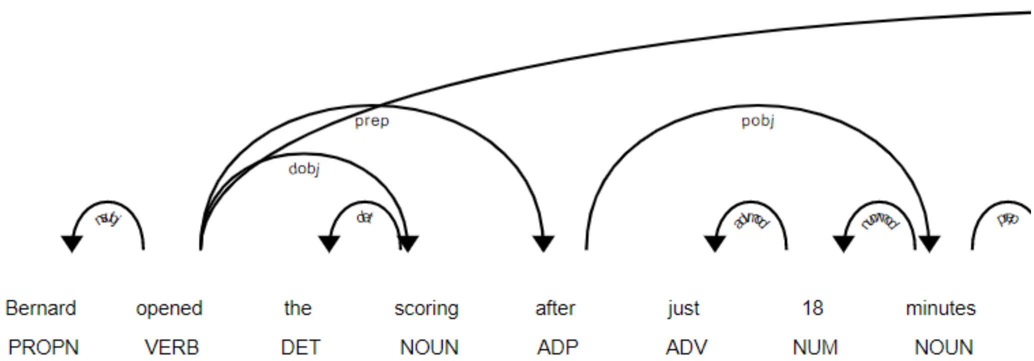


Figure 51. The Visualisation of dependencies

From the vocabulary relationship diagram above, we first determine that the verb "opened" is an action word, and then, based on the action word, we use the lexical relationship "nsubj" to identify the sender as "Bernard." Then, we use the lexical relationship "dobj" to locate the object word "scoring" that corresponds to the action word, where the object signifies the intended action outcome. The system determines the event's initiator and outcome based on an examination of action words and vocabulary dependencies. Consequently, the system uses the team names corresponding to the issuer in the dictionary to generate score statistics for the relevant teams.

Next, we need to identify when the event occurred. First, we use the "minutes" phrase to determine if the event has a specific time. If there is a linked time vocabulary through the "nummod" dependency relationship, the system determines whether the associated quantity in the text is related to the event's time. Eventually, the system can ascertain when the event occurred. Therefore, through the above analysis, we can identify the events contained in the text as:

Bernard scored a goal for Everton at 18 minutes.

Next, the system will extract the appropriate images from the image library in order to generate the images relating to the event. As shown below:

Figure 52. The result for the first sentence

Through the same principle, visualise all the events contained in the text and finally combine them into a comic, as shown in Figure 53.

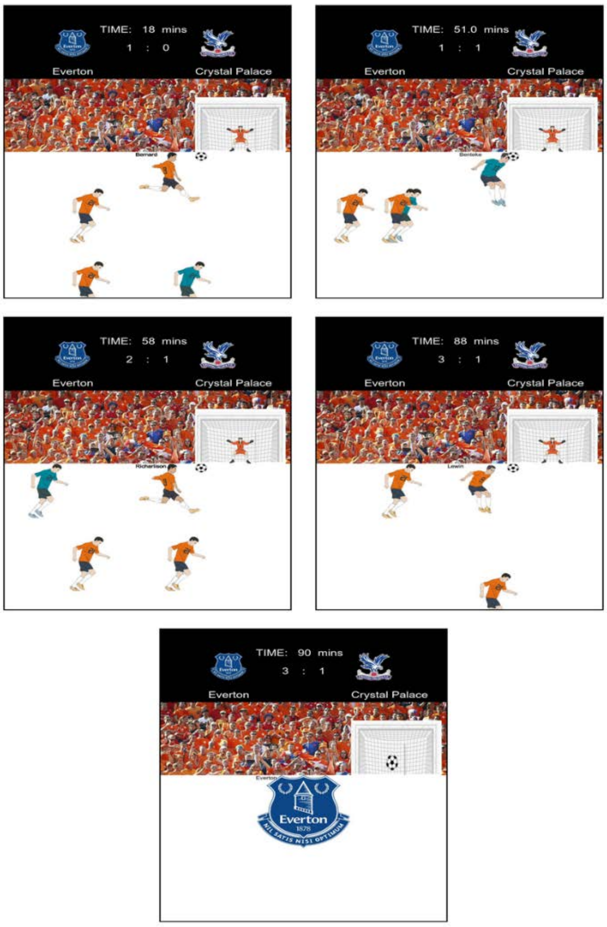


Figure 53. The final result for comic

The comic system's output pictures described above are overly simplistic and monotonous, failing to depict all the participants and the shifting venues in the soccer game. As a result, I added a new background image to the image library during this process and modified the system so that the participating teams attacked in opposite directions in the first and second halves.

Figure 54. The vivid comic for report

As the ball should be in the goal after a goal, I fine-tuned the final image to ensure its accurate representation of the goal. Simultaneously, I adjusted the visual angles to increase the comic's diversity.

Figure 55. The new comic for report

4 Web Experiment

In the initial experiment, we sought to determine whether people gain the same information when reading sports reports presented in the original (long) format, a text summary, or a comic generated by our system. We conducted a web-based study, in which we invited participants to answer questions about the three different media. We concentrated our study on the comprehension effect of the medium.

4.1 The Design of Web Experiment

When we read a story, we concern ourselves with character, time, place, events, and outcome. Thus, the story's analysis differs from the analysis of the sentence. Theorists propose that rules or a 'grammar' exist to help the reader understand the story (Mandler and Johnson, 1977), although Bereiter (1983) offers a critique of this concept. This suggests that, regardless of the type of story being read, the reader's initial reaction is to seek information about the character, time, place, and events in the story. Text summaries and comics generated using the story grammar will encapsulate all the critical information about the event. So, does a difference in readability and helpfulness to readers exist between the original text, the text intermediary, and comics? We evaluated whether an impact on the material's readability existed by comparing the amount of time readers spent reading the three types of materials from different stories and the correct rate/agreement rate of question responses.

In this experiment, we used experimental materials all sourced from real internet soccer game reports to ensure the experiment's authenticity. We prepared 18 sets of questionnaires for three different soccer games, each featuring different types of material from various games. Each participant answered only one set of questionnaires, which also mitigated the influence of the same material's short-term memory on the results. We distributed questionnaires online, and participants chose whether or not to participate in the experimental survey after reading the experimental statement. Participants joined the experiment anonymously throughout, and we only gathered statistics of personal information such as mother tongue, education level, gender, and interest in football. We used a balanced Latin Squares design to distribute the material, ensuring each combination occurred with equal frequency. We allowed two different participants to conduct experiments with the same questionnaire, forming a comparison of the experimental results and averting the interference of subjective participant factors. We paid attention to each participant's reading time, response time, correct answer rate, and subjective question agreement rate during the experimental investigation. By analysing these factors, we determined which type of material best aided readers in understanding the material content. We found that the highest scores were for comic materials by comparing the scores of FRDs(table 3) between different materials, indicating that comic materials more significantly aided participants in understanding events. The scores for the original text and the summary were comparable, but the summary provided less information. So, we assume that readers spent the least amount of time on the comic material and had the highest response and synergy rates, followed by the summary material, with the original text performing the worst.

4.1.1 Participants

We recruited participants through invitations on social media platforms such as WeChat and Facebook. If someone responded to the invitation, we sent them a link to the questionnaire administered through the Questionnaire Star system. We collected no demographic data, so we cannot comment on participants' age, nationality, or interest in soccer. We allocated participants to 1 of 18 combinations (Story x Media) based on appearance. We recorded data only after a person had answered all the questions. Some participants began the experiment but did not complete it. In two other cases, participants appeared to have chosen option A for all questions. We did not interpret this as a serious attempt and discarded them. In this case, we reassigned the combination to the next participant. We allowed participation until all 18 combinations had repeated twice, which means that our analysis relies on the complete responses from 36 participants.

4.1.2 Materials

For this experiment, we have chosen to cover three soccer games Everton vs Crystal Palace (8th February 2020), Bournemouth vs Crawley (26th January 2021), and Newcastle vs Crystal Palace (2nd February 2021). All three games are from the Sky Sports website[2][3][4]. n the system design chapter, we described how to extract the original text using python-based web information crawling technology and how to extract summaries using a summary extraction system based on the story grammar. We observed that the original and summary texts of the three stories had similar lengths (one page for the full text and around 6 sentences for the summary). Also, we found that the number of events contained in the three stories was similar (around 4 events) after reading the original text. Hence, we selected these three soccer games as experimental materials and transformed their summary texts into corresponding comic pictures using the established comic system.

The understanding of the text by the reader depends on the amount and complexity of the information in the material. Since the three stories originate from different sports reports, the original information in the three stories will also differ. This difference will cause variations in the amount of information in the summaries and comics of the three stories. As a result, differences may occur in reading time and total time between the stories or the medium used.

To investigate this possibility, we compared the estimated readability of the stories in three formats (original, summary, and comic) using the Flesch Reading Ease Score (FRES). We calculate this score using a formula, which indicates the reading level (in

[2]https://www.skysports.com/football/bmouth-vs-crawley/report/440591
[3]https://www.skysports.com/football/everton-vs-c-palace/408235
[4]https://www.skysports.com/football/newcastle-united-vs-crystal-palace/429055

terms of US School grade) suitable for the text. We found few examples of FRES applied to comics, so we made some assumptions to facilitate our calculations. We counted the number of words and sentences in the text. Then, we input the full text into the syllable count website (https://syllablecounter.net/count) for syllable statistics. Finally, we used the formula:

206.835 – 1.015 (total words / total sentences) – 84.6 (total syllables / total words)

This formula is used to calculate the score for the corresponding material. One approach is to count the text used in the comic (which would provide a 'fair' comparison with other materials). However, this only counts 'words' in terms of textual entities. We might assume that the reader would fixate on certain visual elements in each panel, such as the crest of each team, the ball, the goal, the larger player depicted as scoring a goal, etc., which might also represent a 'single syllable word'.

Table 3. Flesch Reading Ease score for 3 storys

Story	Original	Summary	Comic
1	884 words	208 words	51 words
	1270 syllables	328 syllables	61 syllables
	34 sentences	6 sentences	5 sentences
	FRES = 58.9	FRES =38.2	FRES =95.3
	DIFFICULT	**DIFFICULT**	**EASY**
2	548 words	92 words	42 words
	768 syllables	130 syllables	52 syllables
	28 sentences	3 sentences	4 sentences
	FRES =68.4	FRES =56.2	FRES =91.4
	EASY	**DIFFICULT**	**EASY**
3	1343 words	168 words	45 words
	1907 syllables	245 syllables	60 syllables
	65 sentences	9 sentences	4 sentences
	FRES = 65.7	FRES = 64.5	FRES = 82.6
	EASY	**EASY**	**EASY**

Table 3 presents the differences in FRES between stories and media. While we could try to balance scores so that FRES is as similar as possible (e.g., by selecting a large sample of original stories, calculating FRES for each story, and selecting a set with similar scores), this might result in a set that does not reflect the typical variation in sports reports. We noted that original stories ranged from 54 to 67. Scores between 50 and 60 are defined as 'fairly difficult to read', scores between 60 and 70 are defined as 'easily understood by 13- to 15-year-old students', and scores between 70 and 80 are 'fairly easy to read'. From this classification, we might expect differences in effort (as measured by reading time) and accuracy (as measured by responses to questions).

As section 1 points out, comprehension can be literal, inferential, or evaluative. Literal comprehension involves the ability to extract and recall surface-level details; inferential comprehension involves the ability to relate content to prior knowledge; evaluative comprehension requires the reader to reason 'beyond' the text. We created 11 questions adaptable to the three stories. We aimed to make these questions both applicable to the three stories and specific to each individual story. For instance, some of the questions we used for story 2 were:

> *Literal - which player scored first in the game? What time was the first goal scored? which team won the match?*
> *Inferential - what will happen to Bournemouth in the FA Cup competition? what will happen to Crawley Town in the FA Cup competition?*
> *Evaluative - how well did Bournemouth press? how well did Crawley Town defend?*

Using the FRES data and classifications, we can order the material as shown in table 4. Rather than merely assuming that the Comic would yield the best results and the original would yield the worst results, the FRES data suggest a more nuanced pattern of predictions.

Table 4. Material sorting and classification based on FRES

Calculated FRES	Definition	Materials
50 - 60	fairly difficult to read	O1 (54.1); S1 (36.6); C1 (56.9); S2 (55.2)
60- 70	easily understood by 13 to 15year-old students	O2 (67.2); O3 (64.4); S3(60.9); C3(60.1)
70 -80	fairly easy to read	C2(76.8)

4.1.3 How was the web survey created?

Due to the COVID-19 pandemic, large-scale face-to-face experiments are impractical. Also, the experiment, which requires reading materials and answering questions, would consume a significant amount of the participants' time. Therefore, we would implement an online experiment survey with a higher degree of freedom as the most feasible method.

Firstly, since we used check-ins from three events, each set of questionnaires contains different types of material, from different materials, prompting us to use the Latin square method. The Latin square design has shown to generate the least amount of experimental error compared to the randomised permutation and randomised block designs. The Latin grid design and the related Graeco-Latin grid and Hyper-Graeco-Latin grid designs are comparative designs we use to control for row- and column-related variations in field experiments. Note that in the Latin square design, the rows and columns are orthogonal to the treatment. For a well-designed experiment, we recommend careful handling (Fink, 2003):

1. Unbiased (same as possible environmental conditions): Allows accurate and valid comparisons between treatment groups.

2. High precision (uniform selection of experimental materials; increased number of observations): allows for identifying any real effects.

3. Wide range of applications: allow to explore other interference factors.

4. Simple experiments (as simple as possible).

Therefore, we designed 18 groups of questionnaires through the questionnaire design website and set the timing function for each page in the questionnaire. We converted the experimental declaration into an HTML file. We put the connections for the 18 groups of questionnaires in a csv-type table. Finally, we used the Python language and the Flask toolkit to design and deploy the questionnaires on the server. This enabled users to click on the questionnaire link to access the experiment declaration page and then choose to agree to enter the questionnaire. We recycled 18 sets of questionnaires to ensure different participants accessed different questionnaires. We posted links to the questionnaires on various social media sites, such as WeChat, Facebook, etc.

4.1.4 How was the questionnaire designed?

Before designing the questionnaire, we needed to ensure that we minimized the interference of participants' short-term memory on the results of the experiment (Cowan, 2008). We used three different types of material (original text, summary, comics) from three different soccer reports in each set of questionnaires. At the same time, to avoid interference caused by reading fatigue (Jeong, 2012), we designed a total of 18 groups of questionnaires. Each questionnaire includes three different types of materials from three different soccer game reports. We ordered the three materials differently in each set of questionnaires.

We designed two types of questions for the questionnaire - the first being objective questions and the second being subjective questions. Participants can find the answers to the objective questions directly from the material. The subjective questions are questions of reasoning and subjective judgment, where participants need to read the material and then answer the questions based on their own understanding of the material.

4.1.5 Procedure

The University of Birmingham Ethics Committee (ERN_21-1092) granted ethical approval for the experiment. We recorded data only once participants had completed all the trials in the experiment. After accessing the link, participants first read the study description to ensure they understood the content and the nature of the experiment. Once they agreed to conduct the experiment, they were directed to the questionnaire page.

During the experiment, participants would first read the material and then move to the page to answer the questions. For each story, participants read the story in one medium and then answered the questions. After this, they received a different story in a different medium, and answered the questions. All participants read three stories, each in a different medium, e.g., Story 1 in original format (1O), followed by Story 2 in Summary format (2S) and then Story 3 in comic format (3C). The combination of the story with the medium produces 18 possible sequences for the experiment. During the questionnaire experiment phase, we experimented each set of questionnaires between two different participants. From this, the combination will repeat twice for 36 participants. In this way, we can make data comparisons, and the objectivity of the experimental results is guaranteed.

4.1.6 Dependent Variables

Time:

a) Total time

Total time participants spent reading and answering questions on a material

b) Reading time

Participant's reading time on a material

c) Time to answer the question

Time for participants to answer questions about a material

Comprehension:

d) Correct answer

The correct rate of participants answering objective questions about a material

e) Agreement between participants

Participant's rate of agreement on subjective questions about a material

Attitude:

f) Media Ratings

Participants' evaluation of material properties, such as:

- Participant rating of reading difficulty of each medium,
- Participant rating of information content of presented by each medium,
- Participant rating of whether they would recommend the medium to friends.

4.2 Data Collection and Analysis

4.2.1 How were the data collected and prepared for analysis?

The ***questionnaire star*** website[5] exclusively collected the experimental data. We set the pages of the questionnaire in English to avoid interference with non-native Chinese speakers. We recorded participants only after they had fully completed an experiment. We carried out preliminary manual processing of the collected experimental data. We classified the results obtained for each group of questionnaires in Excel. The rows represent the different parameter values in the questionnaire, and the columns represent the time results obtained and the questionnaire answers. After processing, we obtained the table shown in figure 56 to facilitate our further analysis and processing.

	Turn	Media	Story	Total time(s)	Reading time(ms)	Answering time(ms)	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	R1	R2	R3
Participant ID																				
O1S2C3	1	O1	Everton v Crystal Palace	315.69	237084	78606	1	2	2	2	2	2	1	2	2	3	2	2	2	3
	2	S2	Bournemouth v Crawley	155.108	50919	104189	1	2	2	2	1	3	1	3	2	3	1	1	2	2
	3	C3	Newcastle v Crystal Palace	94.368	40188	54180	3	1	3	1	1	2	2	1	3	1	2	1	1	1
				565.166	328.191															
O1S2C3	1	O1	Everton v Crystal Palace	300.452	217887	82565	3	2	3	1	1	1	3	1	2	1	3	3	1	2
	2	S2	Bournemouth v Crawley	159.868	37583	122285	1	2	3	1	2	1	2	1	2	1	2	1	1	1
	3	C3	Newcastle v Crystal Palace	92.516	12047	80469	3	1	3	1	1	2	2	1	2	1	2	1	1	1
				552.836	267.517															
O1C3S2	1	O1	Everton v Crystal Palace	176.277	106671	69606	1	2	2	1	1	1	3	1	2	1	2	3	1	3
	2	C3	Newcastle v Crystal Palace	91.476	38836	52640	3	1	3	1	1	2	2	3	2	1	2	1	1	1
	3	S2	Bournemouth v Crawley	95.723	39460	56263	1	1	3	1	2	1	2	1	3	1	3	2	2	3
				363.476	184.967															
O1C3S2	1	O1	Everton v Crystal Palace	419.083	384949	34134	1	2	2	1	2	1	3	1	1	1	2	2	1	2
	2	C3	Newcastle v Crystal Palace	183.007	46088	136919	3	1	3	1	1	2	2	3	1	1	2	1	1	1
	3	S2	Bournemouth v Crawley	138.066	73734	64332	1	2	3	1	2	1	3	1	3	1	2	2	2	2
				740.156	504.771															

Figure 56. New data layout after manual processing

[5]https://www.wjx.cn/

4.2.2 What statistical tests were applied, why, and what software was used?

If the experimental data conform to a normal distribution, we would perform an analysis of variance and T-test based on the different results obtained. If the experimental data do not conform to a normal distribution, we will use a non-parametric analysis and a t-test to obtain the differences in time, accuracy, and other aspects of different types of medias. At the same time, in combination with the comparison of means, we can judge which type of material is more readable. We will perform these tests on SPSS statistical software.

4.2.3 Results

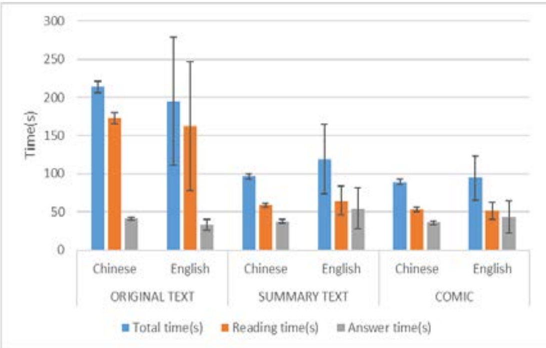
4.2.3.1 Variance Analysis Based on Participant Information

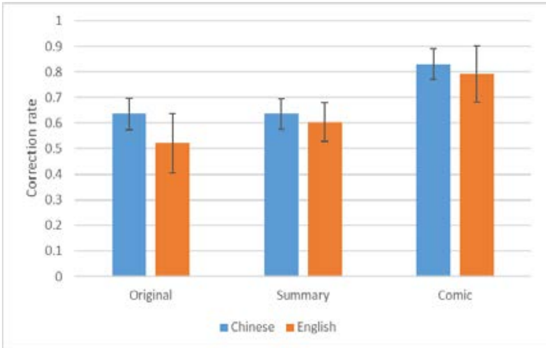
Participants had different background information, so to measure whether these differences in personal information made a significant difference in the results, we grouped participants by first language, gender, interest in football, and current level of education attained.

Table 5. The participant demographic information

Participant information	The amount of participants
Language	
-English	3
-Chinese	69
Gender	
-Male	43
-Female	29
Interest	
-Interested in football	28
-Not interested in football	44
Education	
-A level	20
-Bachelor	21
-Master	31

Due to the very unbalanced information in the groupings based on personal information, we could not perform any statistical analysis on these data. But we can assess possible confounding effects of grouping variables by comparing the means of each indicator. So we made bar graphs of different dependent variables for comparison: English/Chinese; male/female; like soccer/dislike soccer and different educational backgrounds. Since we wrote the material in English and all participants spoke English as a second language in addition to their native language, we used a reference group of 3 English speakers to determine the reading and comprehension baseline scores. The comparison of the participants whose native language is English and Chinese is shown in the figure 57.

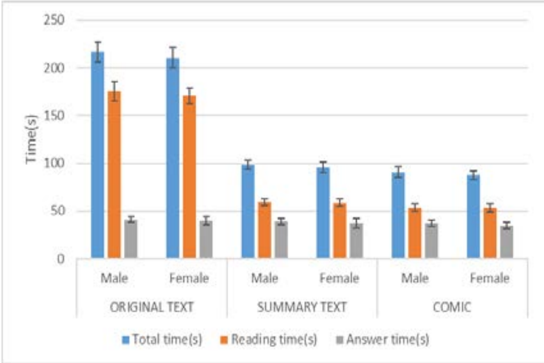




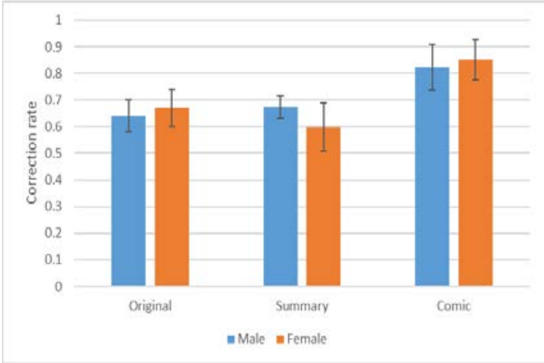
(a) Time distribution based on language (b) Correct distribution based on language

Figure 57. Histogram analysis results of language

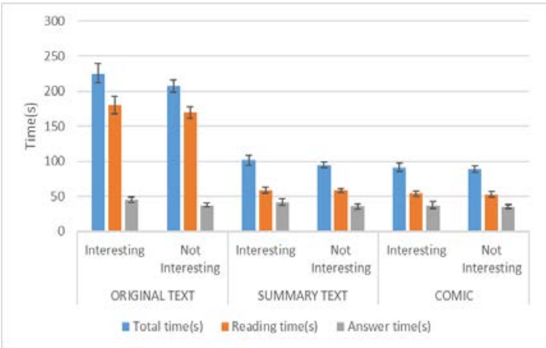
Histograms and overlap of standard deviation intervals show that there are no significant differences in timing or answering questions between the two groups based on different language backgrounds. We also performed a similar analysis for gender, interest in football, and different levels of education. We present its histogram and standard deviation intervals in the figure below.



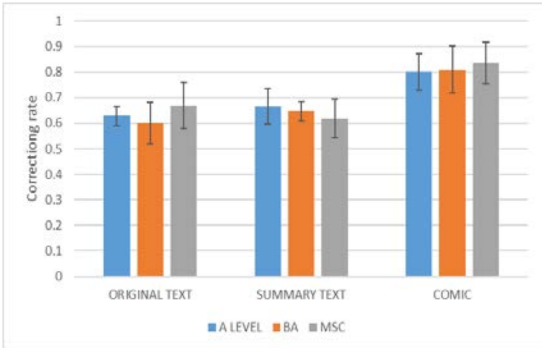
(a) Time distribution based on gender



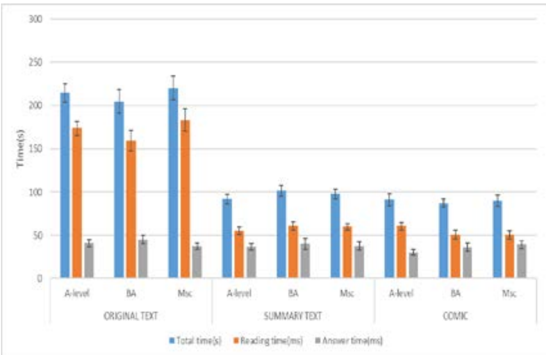
(b) Correct distribution based on gender



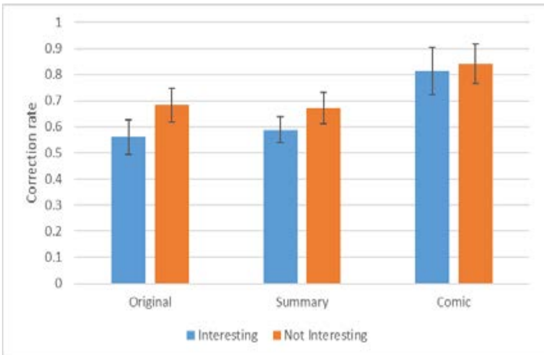
(c) Time distribution based on interest



(d) Correct distribution based on interest



(e) Time distribution based on education



(f) Correct distribution based on education

Figure 58. Histogram analysis results of gender, interes and education

A comparison of the data from the three groups again emphasizes that there were no significant differences between the groups. This suggests that the background factors of the participants did not differ enough to cause a significant difference in the results.

However, through the histogram, we can find that the difference in materials will cause significant differences. Therefore, we can assume that the independent variable of media (original x summary text x cartoon) is sufficient to explain the difference in the dependent variable.

4.2.3.2 Time

g) Total Time

Figure 59 shows the average of the total time for all medias and suggests that time taken to the original version of the story were longer than those of the Summary or Comic versions.

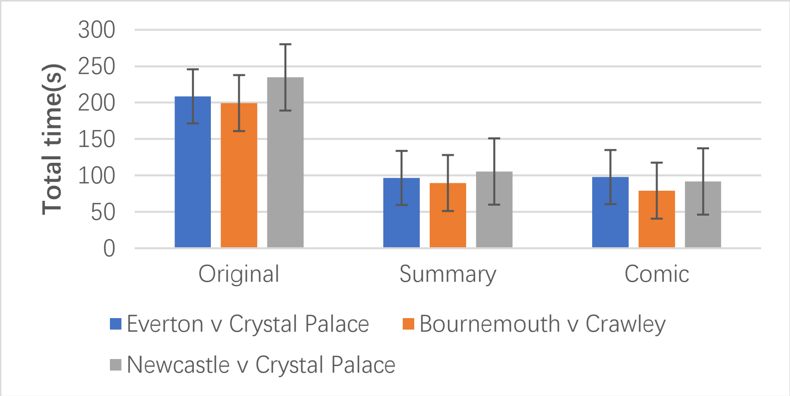


Figure 59. Histogram analysis results of average total time

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
TotalTime	.161	216	.000	.883	216	.000

a. Lilliefors Significance Correction

Figure 60. Normally distribution test based on Shapiro-Wilk for Total Time

From normal distribution, Shapiro-Wilk should have a significance > 0.05. As this is well below the required value, it is safe to assume the data are not normally distributed. This means that we should analyse the data using the Friedman ANOVA test.

Since the data did not conform to a normal distribution, we used a Friedman ANOVA test to analyse the differences in the data. Through the test, we can find that there is a significant main effect of Medium and Story [$x^2(8)$ =125.278, $p<0.05$]. Thus, we can test for significant differences between the conditions. Next, we performed the Wilcoxon test to determine which data are significantly different.

Post-hoc pairwise tests, using the Wilcoxon test, were performed to explore differences between conditions further. We found significant differences between the original text and summary for all stories and between the original text and comics for all stories. But there are occasional non-significant differences between summaries and comics, summaries and summaries, and comics and comics (S1S vs S2C; S3S vs S3C; S1C vs S2C;S2C vs S3C).

h) Reading Time

Figure 61 shows the average for total times across all conditions and suggests that responses to the original version of the story were longer than those to the Summary or Comic versions.

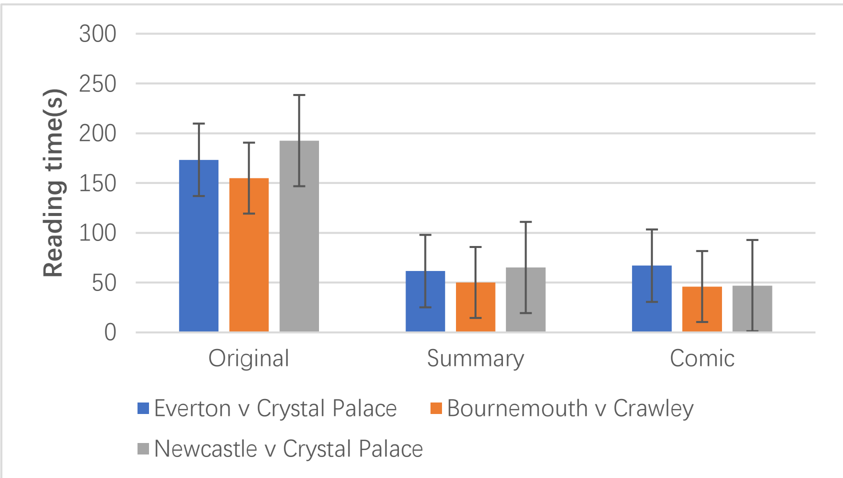


Figure 61. Histogram analysis results of average reading time

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
ReadingTime	.187	216	.000	.859	216	.000

a. Lilliefors Significance Correction

Figure 62. Normally distribution test based on Shapiro-Wilk for Reading Time

From the normal distribution test based on Shapiro-Wilk, which should have significance > 0.05. As this is much (Sig<0.05) less than the required value, it is safe to assume that the data are not normally distributed. This means that we should analyse the data using the Friedman ANOVA test.

Friedman Analysis of Variance, with media and story as variables, showed a main effect [x2 (8) = 136.8, p<0.05]. Post-hoc comparisons, using Mann-Whitney tests, showed significant differences between original and summary or comics but not between summary and comic (as figure 61 shows). While we would expect the original (being much longer than the summary) took longer to read. More interestingly, the comic (even though it has fewer words than the summary) was not read significantly faster than the summary. This suggests that total time is not simply a matter of extracting text but also involves assimilating the information required to answer the questions.

Post-hoc pairwise tests, using the Wilcoxon test, were used to explore differences between conditions further. We found significant differences between original text and summary for all stories and between the original text and comic for all stories. However, there were occasional non-significant differences between original text and original text (S2O vs S1O; S3O vs S1O), summaries and comics (S3S vs S1C; S2S vs S2C; S3S vs S3C), summaries and summaries (S3S vs S1S), comics and comics (S3C vs S2C). A possible explanation for this will be presented in the discussion.

i) Time to Answer Questions

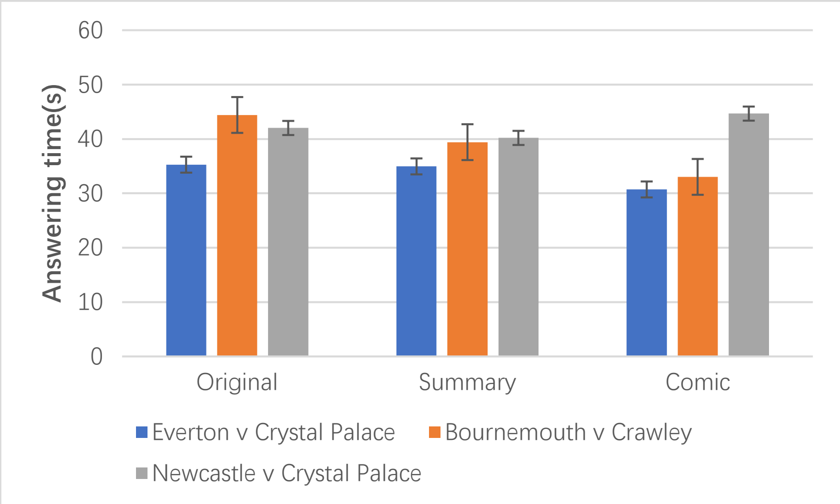


Figure 63. Histogram analysis results of average answering time

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
AnsweringTime	.126	216	.000	.877	216	.000

a. Lilliefors Significance Correction

Figure 64. Normally distribution test based on Shapiro-Wilk for Answering Time

From normally distribution test based on Shapiro-Wilk, which should have significance > 0.05. As this is much (Sig<0.05) less than the required value, it is safe to assume that the data are not normally distributed. This means that we should use a non-parametric test to analyse the data.

Since the data did not conform to a normal distribution, we used Friedman ANOVA to analyse the differences in the data. Through the Friedman ANOVA test, we found that there was a main effect of Medium and Story [x2(8) =17.62, 0.001<p<0.05]. Thus, we concluded there was a difference between the data for answering time. Next, we performed a paired Wilcoxon test to determine which data are significantly different.

Post-hoc pairwise testing, using the Wilcoxon test, was used to explore differences between conditions further. The results revealed only four differences. There were two differences between the original text and comic (S2O vs S1C: $p<0.05$; S2C vs S2O: $p<0.05$). There was only one difference between summary and comic (S3C vs S1S: $p<0.05$) and one difference between comic and comic (S3C vs S1C: $p<0.05$). The results for pairwise comparisons are somewhat different from those obtained from non-parametric tests. A possible explanation for this will be presented in the discussion.

4.2.3.3 Correct and Agreement rate

j) Comprehension

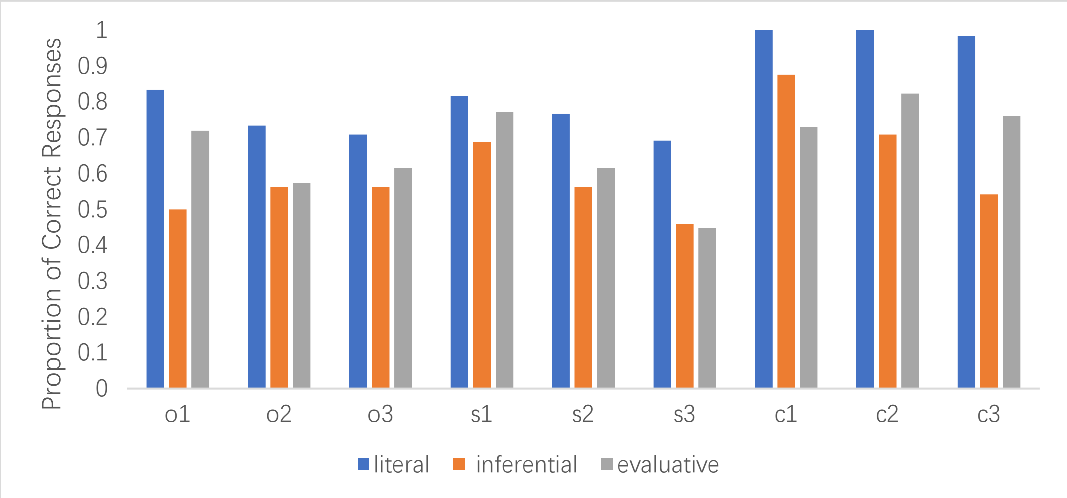


Figure 65. Histogram of correct responses across difference comprehension types (o = original, s = summary, c = comic, and numbers indicate stories 1-3)

Figure 65 suggests differences between comprehension type, media and stories. As our primary interest lies in comparing media, we decided to pool the story data for each medium for subsequent analysis. Friedman ANOVA was applied to the comprehension tests (literal x inferential x evaluative) for each medium.

For the original text, there was a significant main effect of comprehension type [$x2(2) = 20.9$, $p<0.001$]. The mean correct response for literal was 0.76; for evaluative, it was 0.6; and for inferential, it was 0.54. Post-hoc, pairwise comparisons, using Wilcoxon tests, showed significant differences between literal and evaluative ($z= 3.6$, $p<0.001$), literal and inferential ($z = 4.4$, $p<0.001$). There was no difference between evaluative and inferential.

In summary, there was also a significant main effect of comprehension type [$x2(2) = 24.3$, $p<0.001$]. The mean correct response for literal was 0.76; for evaluative, it was 0.59; and for inferential, it was 0.57. Post-hoc, pairwise comparisons, using Wilcoxon tests, showed significant differences between literal and evaluative ($z=64$, $p<0.001$), literal and inferential ($z = 3.9$, $p<0.001$). There was no difference between evaluative and inferential.

Finally, comics also showed a significant main effect on comprehension type [$x2(2) = 56.2$, $p<0.001$]. The mean correct response for literal was 0.99; for evaluative, it was 0.77; for inferential, it was 0.73. Post-hoc, pairwise comparisons, using Wilcoxon tests, showed significant differences between literal and evaluative ($z=6.1$, $p<0.001$), literal and inferential ($z = 6.1$, $p<0.001$). There was no difference between evaluative and inferential.

It is interesting to note that the mean response rate for comics was higher than the other media. Tests for each type of comprehension test were conducted for each medium.
For literal comprehension, there is a significant main effect of medium [$x2(2) = 63.6$, $p<0.001$]. Post-hoc, pairwise comparisons, using Wilcoxon tests, showed significant differences between the comic and original ($z=6.3$, $p<0.001$) and comic and summary ($z = 5.8$, $p<0.001$), but not between original and summary.

For evaluative comprehension, there was a significant main effect of medium [$x2(2) = 16.7$, $p<0.001$]. Post-hoc, pairwise comparisons, using Wilcoxon tests, showed significant differences between the comic and original ($z=3.8$, $p<0.001$) and comic and summary ($z = 3.9$, $p<0.001$), but not between original and summary.

For inferential comprehension, there was a significant main effect of medium [$x2(2) = 11.7$, $p<0.05$]. Post-hoc, pairwise comparisons, using Wilcoxon tests, showed significant differences between the comic and original (z=3, $p<0.05$) and comic and summary (z = 3.4, $p<0.001$), but not between original and summary.

k) Agreement rate

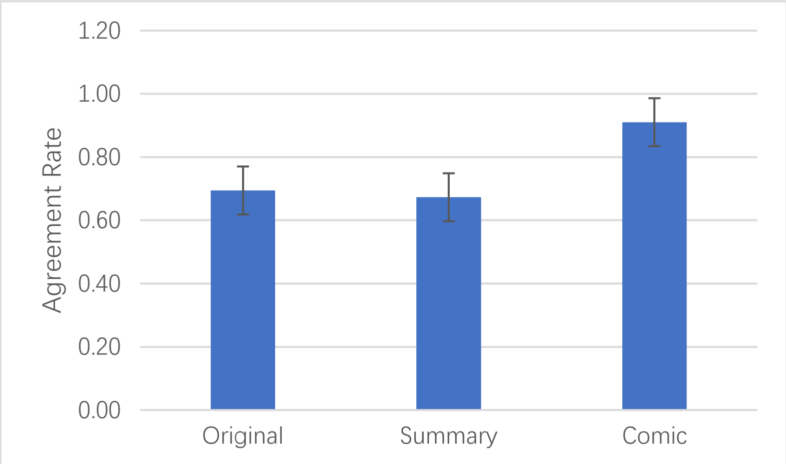


Figure 66. Histogram analysis results of average agreement rate

From Figure 66, we see that the agreement rate of comics is much higher than that of summary and original text. This also means that more readers have a higher degree of uniformity in choosing answers when answering questions.

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Aggrement	.186	99	.000	.879	99	.000

a. Lilliefors Significance Correction

Figure 67. Normally distribution test based on Shapiro-Wilk for Agreement Rate

From normally distribution test based on Shapiro-Wilk, which should have significance > 0.05. As this is much (Sig<0.05) less than the required value, it is safe to assume that the results are not normally distributed. This means that we should analyse the data using the Friedman ANOVA tests.

Since the data did not conform to a normal distribution, we used Friedman ANOVA tests to analyse the differences in the data. Through the test, we found that there was a main effect of Medium and Story [x2(8) =54.451, p<0.05]. Thus, we concluded that there was a significant difference between the data for total time. Next, we performed the Wilcoxon test to determine which data are significantly different.

Post-hoc pairwise testing was conducted using the Wilcoxon test to explore differences between conditions further. Results found significant differences between Original and Comic for all stories and between summary and Comic for all stories. There was one significant difference between the original text and the summary (S1O vs S3S: p<0.05). And There was one significant difference between the summary and the summary (S1S vs S3S: p<0.05).

l) Rating of Media

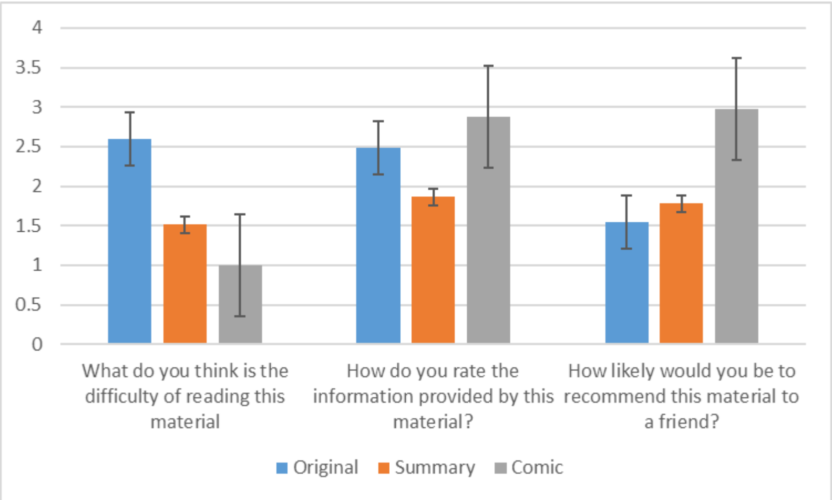


Figure 68. Histogram analysis results of Rating of media

Figure 68 shows the average rating of media across all conditions and shows that responding to the story in three types of material is very different. We will do a two-factor ANOVA to check the P-value.

Question1- What do you think is the difficulty of reading this material?

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
R1	.344	216	.000	.719	216	.000

a. Lilliefors Significance Correction

Figure 69. Normally distribution test based on Shapiro-Wilk for Question 1

From normally distribution test based on Shapiro-Wilk, which should have significance > 0.05. As this is much (Sig<0.05) less than the required value, it is safe to assume that the data are not normally distributed. This means that we should analyse the data using the Friedman ANOVA tests.

Since the data did not conform to a normal distribution, we used Friedman ANOVA tests to analyse the differences in the data. Through the Friedman ANOVA test, we found that there is a main effect of Medium and Story [x2(8) =132.435, p<0.05]. Thus, we concluded that there was a significant difference between the data for the total time. Next, we performed the Wilcoxon test to determine which data are significantly different.

Post-hoc pairwise testing, using the Wilcoxon test, was used to explore differences between conditions further. We found that for the first media evaluation question, there were significant differences between different types of material. There is no significant difference between the same type of materials (except S2S vs S3S: p<0.05).

Question2- How do you rate the information provided by this material?

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
R2	.340	216	.000	.734	216	.000

a. Lilliefors Significance Correction

Figure 70. Normally distribution test based on Shapiro-Wilk for Question 2

From normally distribution test based on Shapiro-Wilk, which should have significance > 0.05. As this is much (Sig<0.05) less than the required value, it is safe to assume that the data are not normally distributed. This means that we should analyse the data using the Friedman ANOVA tests.

Since the data did not conform to a normal distribution, we used Friedman ANOVA tests to analyse the differences in the data. Through the Friedman ANOVA test, we found that there was a main effect of Medium and Story [x2(8) =75.904, p<0.05]. Thus, we concluded that there was a significant difference between the data for the total time. Next, we performed the Wilcoxon test to determine which data are significantly different.

Post-hoc pairwise testing, using the Wilcoxon test, was used to explore differences between conditions further. It was found that there was no significant difference between the same type of materials. We found that for the second media evaluation question, there were significant differences between most of the different types of materials. However, there are still some materials that did not differ from each other (S2C vs S1O: p>0.05; S2O vs S1S: p>0.05; S3O vs S1C: p>0.05; S2S vs S2O: p>0.05;).

Question3-How likely would you be to recommend this material to a friend?

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
R3	.262	216	.000	.778	216	.000

a. Lilliefors Significance Correction

Figure 71. Normally distribution test based on Shapiro-Wilk for Question 3

From normally distribution test based on Shapiro-Wilk, which should have significance > 0.05. As this is much (Sig<0.05) less than the required value, it is safe to assume that the data are not normally distributed. This means that we should analyse the data using the

Friedman ANOVA tests.

Since the data did not conform to a normal distribution, we used Friedman ANOVA tests to analyse the differences in the data. Through the test, we found that there was a main effect of Medium and Story [$x2(8) = 122.278$, $p < 0.05$]. Thus, we concluded that there was a significant difference between the data for the total time. Next, we performed the Wilcoxon test to determine which data are significantly different.

Post-hoc pairwise testing, using the Wilcoxon test, was used to explore differences between conditions further. It was found that there were significant differences between Original and Comic for all stories and significant differences between summary and Comic for all stories. There was no significant difference between the original text and the summary.

4.3 Conclusions from Web Experiment

Through the comparison of different dependent variables, we found that the individual factors of the participants were not enough to cause significant differences in the results. However, there are significant differences in the experimental results between different media. This means that the influence of the material on comprehension is the most significant when there is a foundation in English.、

The reading time for the original material was significantly longer than for the summary or comic, as expected due to the original material's length. However, the reading times for summary and comic were similar. This suggested that reading the comic involved more than just fixating on the text contained within. In experiment 2, we plan to explore the reading strategy for both text and comic. We noticed that the reading times for stories 1 and 2 display differences between the comics and the summaries, reflecting partially the differences in their FRES scores (table 3). Regarding answering the questions, we found no differences between the media. This could be due to participants using their recall of the content of the story to answer the questions. The performance was significantly better when reading the comics in terms of answering the questions. There was no significant difference between the original text and the summary text, which surprised us, as the length of the Original text might have overwhelmed participants and led them to make more mistakes than the shorter Summary. In terms of Consistency, participants were less consistent with text (Original or Summary), suggesting that they were selecting wrong answers at random when uncertain. This implies that they formed an incomplete memory of the information in the text they read (and this was not dependent on the text's length).

Concerning comprehension, literal comprehension produced a significantly superior performance to evaluative or inferential. This is unsurprising given that this only required identifying and recalling salient material. Comics also performed well in terms of literal comprehension (because the participants had fewer words to review and little need to filter extraneous content). What surprised us was that the comics also produced superior performance in evaluative and inferential comprehension. As there was no difference between the original text and the summary in terms of evaluative and inferential comprehension, we argue that this cannot be attributed merely to the number of words. Rather, it might be that the text versions (original and summary) created sufficient confusion, through the 'noise' of additional material, to make it more difficult for participants to form a mental model (Johnson-Laird, 1992) sufficient to support comprehension in the way we tested it. However, the comic's visualisation of the material may be sufficient to support an interpretation beyond the text provided. While we did not explore this in the study, it offers the intriguing possibility that the comic summary engages a different form of reading than that supported by the text alone.

Participants were more likely to recommend the comic to a friend and believe that the comic provided salient information. The rating of reading difficulty indicates that participants considered the original text the most difficult to read, while the comic was the easiest. Interestingly, in terms of ease of reading, the Summary was rated close to the

Comic but the worst in providing salient information. This suggests that participants were aware of their performance in answering the questions.

5 Eye-Tracking Experiment

The performance of the question responses in experiment one may have resulted from the original text being challenging to read. This is supported by the participants' subjective responses to the evaluation questions, as the summary did not contain all the material they felt was salient to the questions. This could mean that the text presentation of the information might have impaired their ability to remember the content. However, this does not explain why the comic (which contained the least amount of text) led to better performance in answering the questions. To consider this, we decided to use the eye-tracking method to compare reading strategies across media.

5.1 The Design of Eye-Tracking experiment

Cohn and other researchers have emphasised the need to understand how people ‘read’ comics(Eisner, 2008a; Cohn, 2013b). That is, instead of treating the comic as a visual medium, Cohn(Cohn, 2013b) convincingly argues that readers will organise text and images into well-formed sequences and apply a 'narrative grammar' that uses similar strategies to the processing of words when reading text. Cohn's (2013) theory of reading comics suggests that people construct a 'narrative grammar' that combines text and image to find meaning in the content. He further proposes that (Western) readers might seek to 'read' the comic similarly to text, i.e., using a Z-pattern of eye movement (left-to-right and downward reading order) familiar from reading Western texts. Indeed, when the comic is laid out as a grid, eye movement tends to follow such a pattern.

In addition to the layout of panels affecting reading strategy, eye-tracking can also reveal how content affects the 'entry point' to the material. We know, from studies of reading newspapers, that the size and colour of the information affect where people look first (Holsanova et al., 2006). Such an entry point would result in a longer fixation time, with the reader developing inferences to guide their reading. These inferences will, in turn, depend on the 'task' (Ballard and Hayhoe, 2009). Reading a comic might require following the plot, enjoying the 'action', interpreting the characters' motivation, and remembering details for later recall (Mikkonen and Lautenbacher, 2019; Cohn, 2013b; Cohn, 2020). Capturing eye movement and entry points could be informative in exploring how readers use comics to interpret sports events, particularly in terms of how they judge information to be salient.

In the eye-tracking experiment, we primarily tracked the gaze of participants through an eye-tracking device. With eye-tracking experiments, we can obtain the position of the user's fixation point and the intensity (Pomorska) of the fixation, which reflects the fixed position of the participant's gaze when reading the article/comic. The location and intensity of the fixation point would also reflect the participants' interest in the area. Furthermore, through the movement of the fixed point of sight, we can obtain the movement trajectory of the participant's sight when reading the material. The user's saccade trajectory would reflect the reading habits of the participants and whether there is information in the material that can effectively stimulate vision. For example, spanning a larger saccade can indicate that the participant is looking for meaningful information at a distance, or a high frequency of repetitive saccades can indicate that the participant has a larger camp for the area or that the participants wanted more information through extensive visual search. (Kurzhals et al., 2016;Djamasbi et al., 2011;Buscher et al., 2009). Through the positions and changes of visual fixation points, we were able to obtain participants' reading trajectories in different types of materials and regions of interest.

Simultaneously, we will analyse and compare the collected data to determine which material would be more conducive to readers' understanding. In general, picture information is more specific than text information. Thus, we hypothesised that comics would be the best in terms of readability.

We need to evaluate any text summarisation process by readers in terms of the meaningfulness of the output. There is a possibility (particularly with text extraction) that salient information is omitted or the resulting output is difficult to comprehend. This might be particularly problematic when the comic reduces the original story to so few words. We, therefore, used a test of literal comprehension to compare the ability of readers to understand the content of sports reports. In our study, these questions took the following form:

Which team scored the first goal?
Which team got the last score in the game?
What was the name of the player who scored the first goal?
What was the name of the player who got the last goal in the game?
When was the first goal scored?
What was the final score of the match?
Which team won the game?

We defined the correct answers for each story with seven questions, based on the content of the report. The aim of the eye-tracking experiment is to gather data about the participant's eye movements and the ability of participants to extract information from different types of material. Therefore, the design of the questions doesn't need to be overly complex; it only needs to closely relate to the information from the material. Although the aforementioned seven questions are simple and can be answered directly from the material, they reflect the participants' understanding of the material more intuitively. For consistency of response, we presented the questions as multiple-choice, with three options for each question. Therefore, it is a test of literal comprehension that should be unchallenging and easy to perform. In addition, we will measure reading time, assuming that 'efficiency' could be defined by faster reading and more correct answers.

Moreover, considering the discussion in the introduction about how people read comics, we were interested in whether the comics we generated are consistent with previous research on comic reading.

5.1.1 Participants

The experiment involved 12 participants. None of the participants were soccer enthusiasts. We deemed this appropriate as a screening criterion because we didn't want the results to be influenced by prior knowledge. As much as possible, we wanted participants' answers to the questions to be based on the information they obtained from reading the provided material.

5.1.2 Equipment

We conducted the reading experiment using a screen-mounted Tobii eye tracker (X2-60), running at 60Hz. We used the standard 9-point calibration from the manufacturer. This involves a grid of 9 dots to cover the screen, with the participant being cued to fixate on each dot individually (the dot grows larger to indicate which one to look at). At the end of the calibration, the grid is shown with the fixation points overlaid. If the points are (a) green and (b) closely positioned on or near the dots, then we accept the calibration. Where this is not the case, we repeat the calibration.

5.1.3 Materials

We used 3 different event reports. The full text is a report of a match between Bournemouth and Crawley Town on 26th January 2021[6] ; the summary text is the game report between Everton and Crystal Palace on 8th February 2020[7]; the comic is a report of the match between Newcastle United and Crystal Palace on 2nd February 2021[8].

---

[6] https://www.skysports.com/football/bmouth-vs-crawley/report/440591

[7] https://www.skysports.com/football/everton-vs-c-palace/408235

[8] https://www.skysports.com/football/newcastle-united-vs-crystal-palace/429055

We compared readability estimates for three formats of stories (full text, summary, and comic). We used the Flesch Reading Ease Score (Bereiter, 1983). This can be calculated by the following (Eisner, 2008a) and indicates the reading level (in terms of US School grade) for which the text would be appropriate.

206.835 – 1.015 (total words / total sentences) – 84.6 (total syllables / total words) (Eisner, 2008a)

Table 6: FRES calculations for different media

	Full-text	Summary	Comic
Words	548	208	45
Syllables	758	328	60
Sentences	28	6	4
FRES	68.4	38.2	82.6
US School grade	8^{th} - 9^{th}	College	10^{th} - 12th

We found few examples of FRES applied to comics, leading us to make certain assumptions to support our calculations. One option we considered is to count the text used in the comic, providing a 'fair' comparison with the other materials. However, this only counts 'words' in terms of textual entities. We made the assumption that each panel in our comic represents a 'sentence'. We further hypothesized that the reader fixates on some visual elements in each panel, such as the crest of each team, the ball, the goal, the larger player scoring a goal, etc., which we might also consider as a 'single syllable word'. While the FRES method is not free from criticism and can produce anomalous results (as it doesn't consider the difficulty or meaningfulness of individual words), the analysis suggests that the level of this material should be fairly easy to read. If we include some image content in our calculation, we find, somewhat surprisingly, that the full-text and comic receive similar scores. All the media types can be graded as easy for 8^{th} or 9^{th}-grade students to read (scores around 60-70), with the Summary text graded as more challenging.

Calculating FRES for the comic as text + images gave us a means of evaluating the different media. We had assumed that the comic would have a FRES score corresponding to a much lower School Grade than the other media. The small difference between them is interesting and raises the question of how to score comics for readability.

Another point to note in our study is that the participants were not soccer fans. Consequently, the information presented in the comic might not be 'salient' for them. In other words, if the names of the soccer teams or players (or even the basic rules of the game) mean little to the readers, then this information may not be 'salient'. We also noted that applying FRES highlights the difficulty of calculating the readability of comics when we consider that images may have semantic content and that comics may pose challenges for some readers in terms of understanding visual content.

5.1.4 Procedure

The University Ethics (ERN_21-1092) approved the study. Upon arrival, we provided participants with an informed consent form and told them they could withdraw from the experiment at any time without facing any penalties. We explained the purpose of the experiment as "studying how people read different versions of sports reports". For each participant, we calibrated the Tobii eye-tracking device using the built-in procedure. Once we accepted the participant's calibration, the experiment began. We used Tobii's built-in experiment design software, which created a slideshow in which screen 1 displayed the instructions. Once the participant had read these, pressing the space bar would bring them to the first set of materials. We counter-balanced the presentation of material, i.e., full-text x summary x comic, with all permutations defined in a balanced Latin Square design. For the full text, participants could use the arrow keys to move between the two pages where we located the material. For the summary and comic, we fit the material on a single page. After the participant was satisfied that they had read the material in sufficient depth, pressing the space-bar moved them to the first of the questions (we also randomized these across participants in terms of question order and answer options). Providing an answer to a question moved the participant to the next question. We only accepted the first answer given, and there was no opportunity to change this answer. Once they answered all 7 questions, we instructed them to press the space bar to access the next material. The experiment was self-paced, and participants were free to spend as long as they wished on each material and question.

5.2 Data Collection and Analysis

We report on participants' performance in answering literal comprehension test questions and reading strategy. For the reading strategy, we focused on the average time per word and average fixation per word. We assume that this provides a measure of comparability across materials, though as we will point out in the Discussion, the assumption of comparability might also imply similar reading strategies, which may not be appropriate.

We evaluated the data for normality using the Shapiro-Wilk test. When the data violated normality, we applied non-parametric statistical tests: a Friedman Analysis of Variance across the three types of material and Wilcoxon Signed-Rank tests for post-hoc pairwise comparison.

5.2.1 Results

The results section begins with a summary of performance (in terms of answers to questions and average reading and fixation times across different media) and then considers the reading strategies used by participants for the comic.

Table 7. Result for Full Text

FULL TEXT					
Participant	Total Time	Time per Word	Total Fixations	Time per Fixation	Total Questions Correct
1	221.685	0.40826	540	0.410528	57.14%
2	157.645	0.290322	526	0.299705	85.71%
3	171.457	0.315759	594	0.288648	85.71%
4	214.531	0.395085	399	0.537672	42.86%
5	124.721	0.229689	244	0.511152	57.14%
6	115.582	0.212858	446	0.259152	42.86%
7	127.042	0.233963	510	0.249102	57.14%
8	287.658	0.529757	864	0.332938	42.86%
9	75.521	0.139081	154	0.490396	28.57%
10	180.608	0.332611	614	0.29415	42.86%
11	88.001	0.162064	311	0.282961	57.14%
12	86.057	0.158484	360	0.239047	28.57%

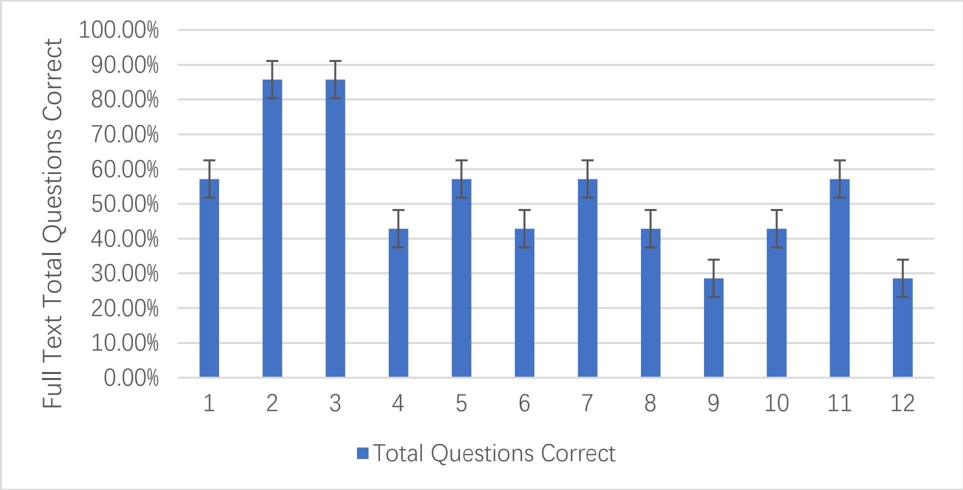


Figure 72. Correct rate histogram for the question corresponding to the original text

Table 8. Result for Summary

SUMMARY					
Participant	Total Time	Time per Word	Total Fixations	Time per Fixation	Total Questions Correct
1	168.891	0.811976	333	0.50718	85.71%
2	127.356	0.612288	221	0.576271	100.00%
3	63.261	0.304139	198	0.3195	28.57%
4	106.624	0.512615	254	0.41978	71.43%
5	83.865	0.403197	215	0.39007	85.71%
6	45.571	0.219091	143	0.318678	57.14%
7	61.071	0.293611	193	0.31643	28.57%
8	214.755	1.032476	638	0.336607	57.14%
9	30.19	0.145144	68	0.443971	57.14%
10	83.907	0.403399	262	0.320256	57.14%
11	49.85	0.239663	148	0.336824	57.14%
12	64.904	0.312038	186	0.348946	71.43%

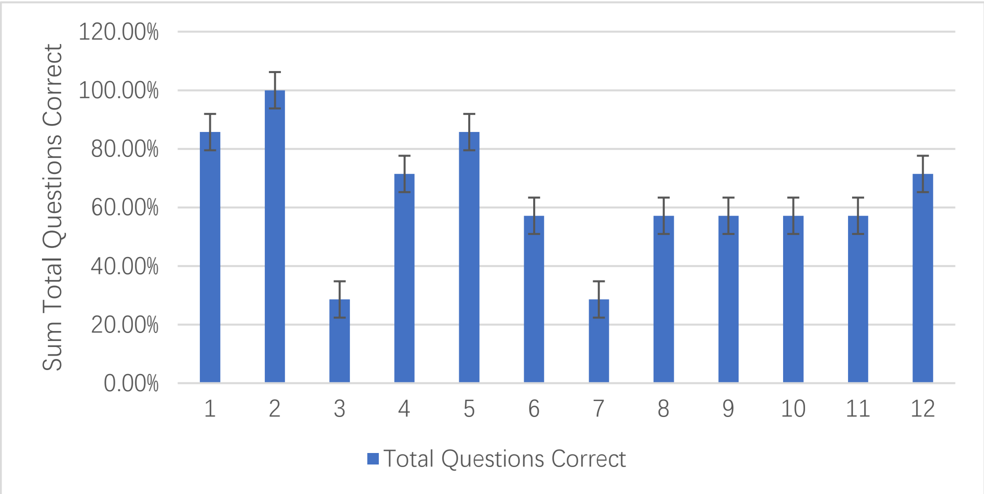


Figure 73. Correct rate histogram for the question corresponding to the Summary text

Table 9. Result for Comic

COMIC					
Participant	Total Time	Time per Word	Total Fixations	Time per Fixation	Total Questions Correct
1	57.25	0.923387	185	0.309459	71.43%
2	56.944	0.918452	92	0.618957	71.43%
3	22.782	0.367452	81	0.281259	71.43%
4	50.82	0.819677	105	0.484	100.00%
5	37.777	0.609306	138	0.273746	71.43%
6	28.343	0.457145	96	0.29524	85.71%
7	81.019	1.306758	54	1.500352	71.43%
8	43.571	0.702758	142	0.306838	100.00%
9	28.345	0.457177	66	0.42947	85.71%
10	25.159	0.40579	88	0.285898	100.00%
11	30.263	0.488113	115	0.263157	85.71%
12	19.755	0.318629	16	1.234688	85.71%

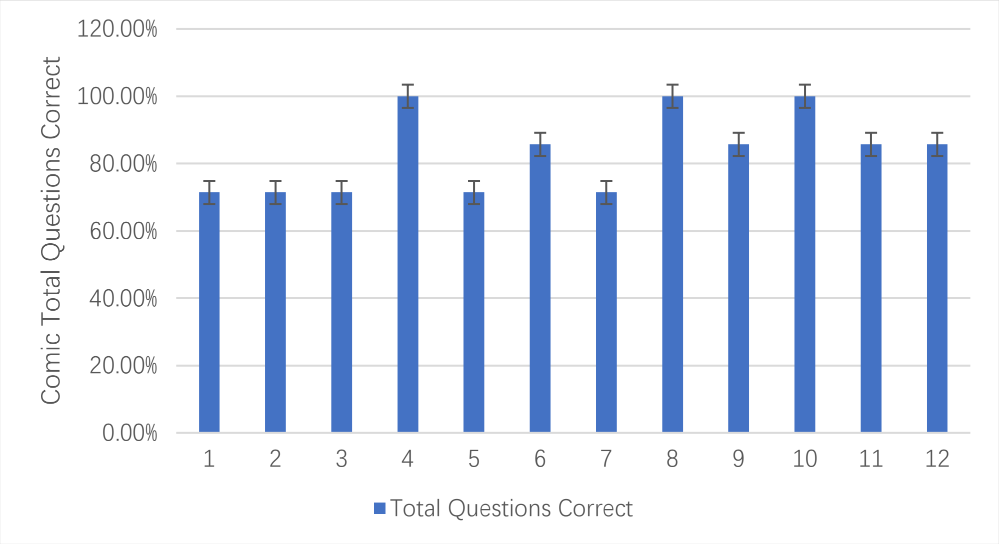


Figure 74. Correct rate histogram for the question corresponding to the comic

m) Answering Questions

Shapiro-Wilk indicated that the data were normally distributed, and consequently, a one-way Analysis of Variance was applied. There was a significant main effect of material on correct answers to question [F(2,22) = 7.941, p<0.005, partial-eta2 = 0.71]. Mauchly's Test of Sphericity indicated that the assumption of sphericity had not been violated, χ2(2) = 0.99, p = 0.94. Subsequent pairwise comparison showed significant difference between comic and full-text [t(11) = 3.8, p<0.005] and comic and summary [t(11) = 2.7, p<0.05], but not between full-text and summary [t(11)= 1.3, p = 0.2]. These differences are illustrated in figure 75.

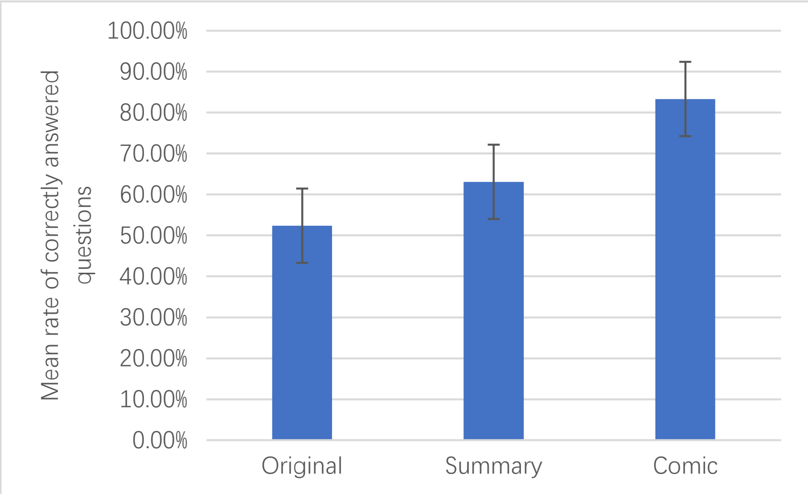


Figure 75: Mean number of correctly answered questions for each condition (CQ = comic questions; FQ = full-text questions; SQ = summary questions)

n) Average Reading Time per Word

There was a significant main effect of material of time per word [$x^2(11) = 23.212$, $p<0.05$].

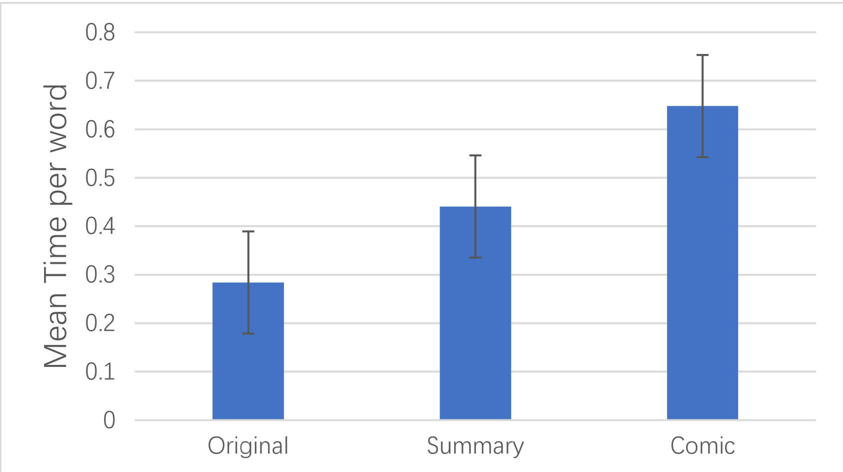


Figure 76: Mean Reading Time per Word for each Condition

Pairwise comparison revealed significant difference between comic and full or summary ($p<0.005$), and between full and summary ($p<0.05$). People spent less time reading each word in the full text and summary conditions than from the comic (figure 76).

o) Average Time per Fixation

There was no effect of material on mean time per fixation [$x^2(11) = 16.246$, $p>0.05$]. This suggests that people were attending to salient items in a similar manner (figure 77). Of course, the fact that the comic had fewer items to which to attend could mean that reading extracts more information.

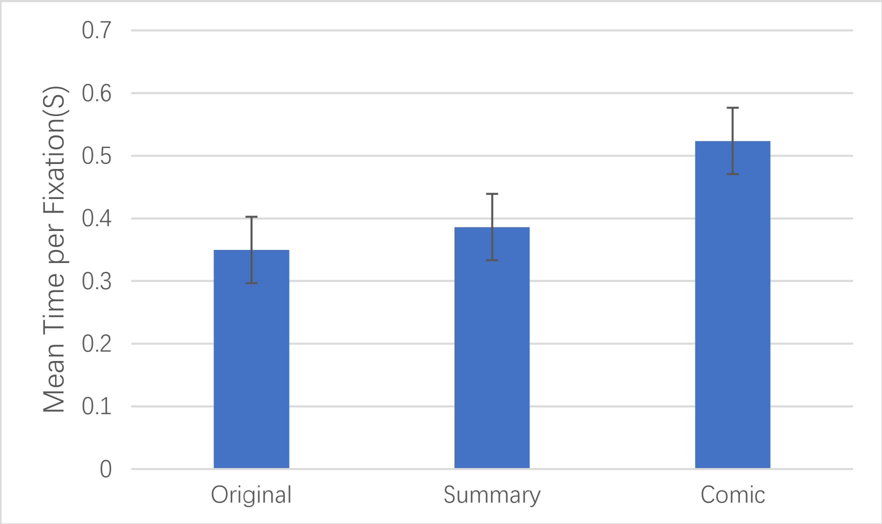


Figure 77: Mean Time per Fixation for each Condition

p) Heat Maps and Reading Strategies for the Comic

As shown in figure 78 and 79, the eye tracking data from the full-text and summary text suggested that participants tended to skim read without fixating on all the words. This could provide a simple explanation for the difference in the ability to answer the questions: in the comic condition, participants only looked at the text and ignored the images. However, as the heat map in figure 80 shows, participants *were* attending to the text and images.

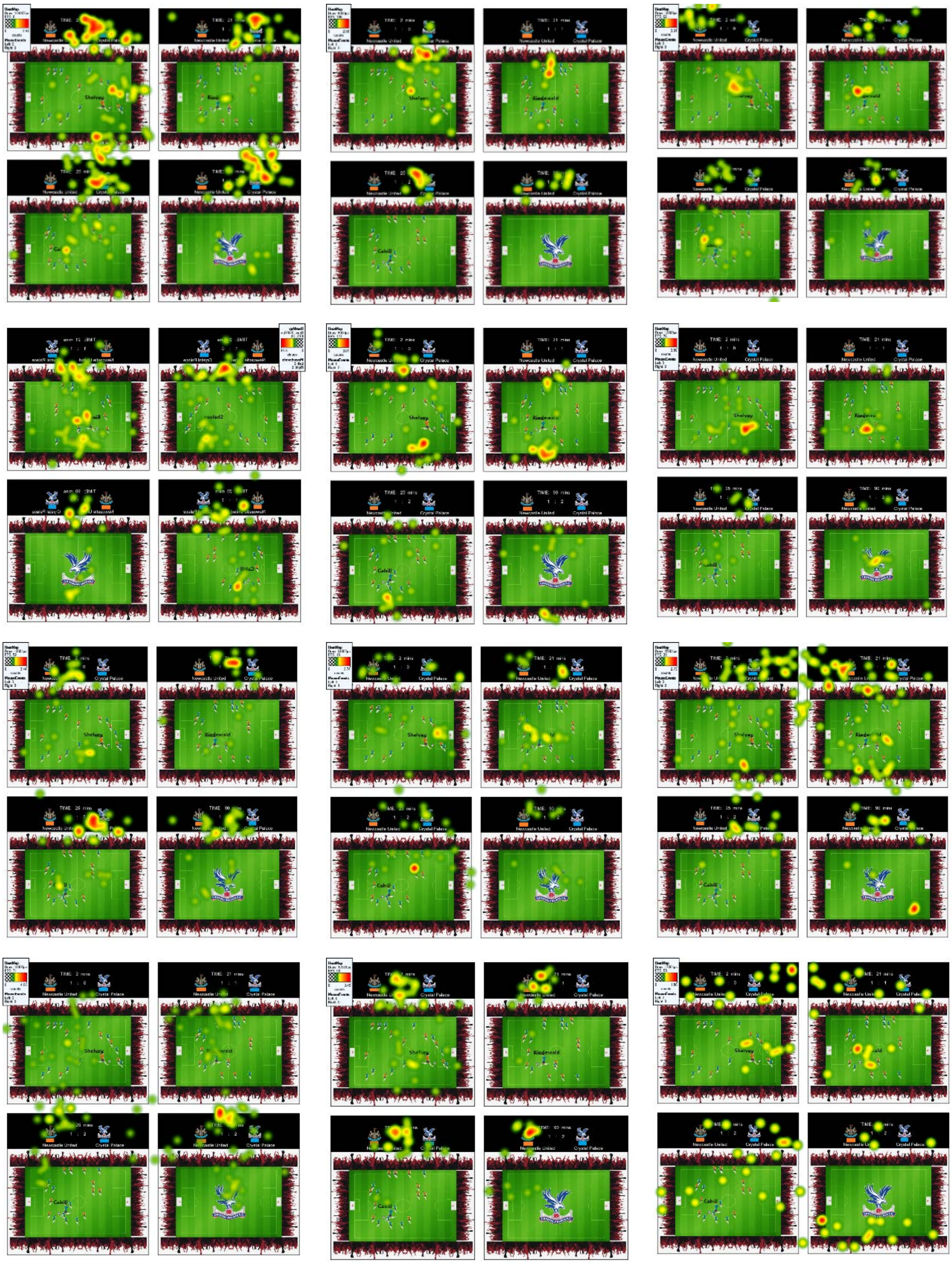


Figure 78: Heat Map for Comic showing pattern of fixations for 12 participants

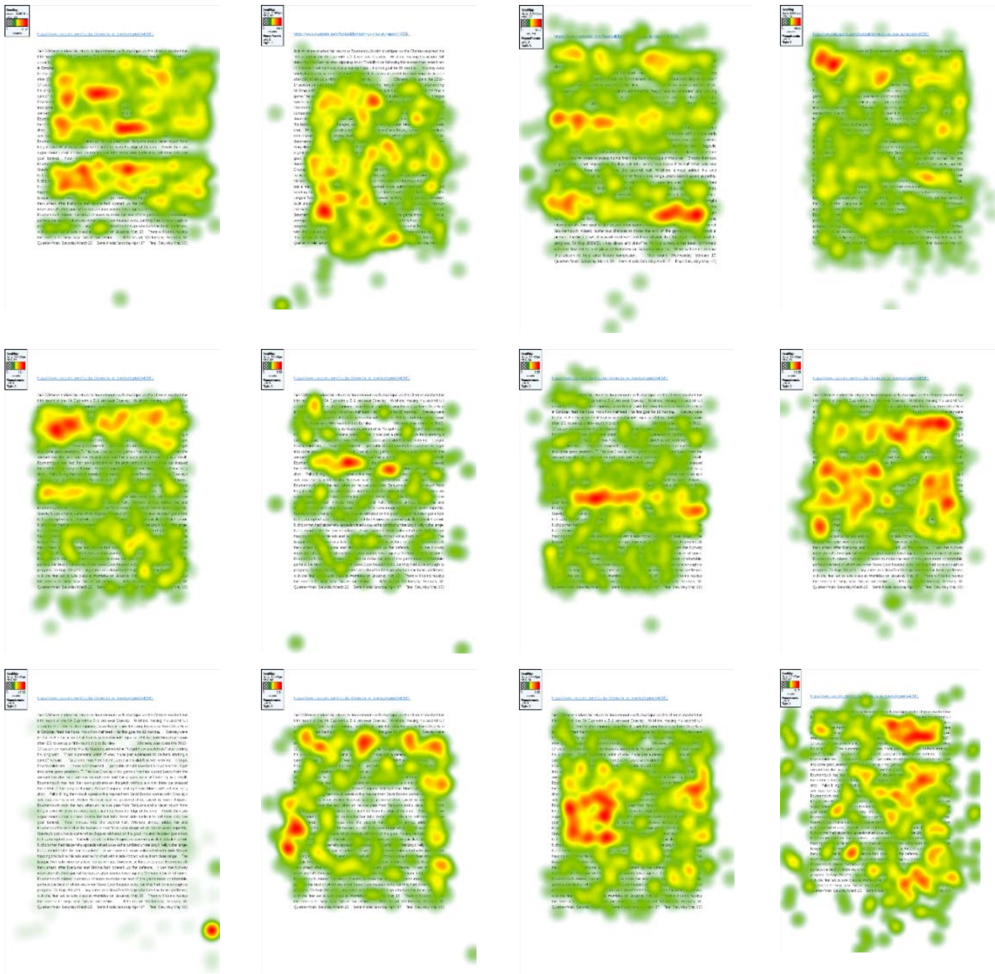


Figure 79: Heat Map for full text showing pattern of fixations for 12 participants

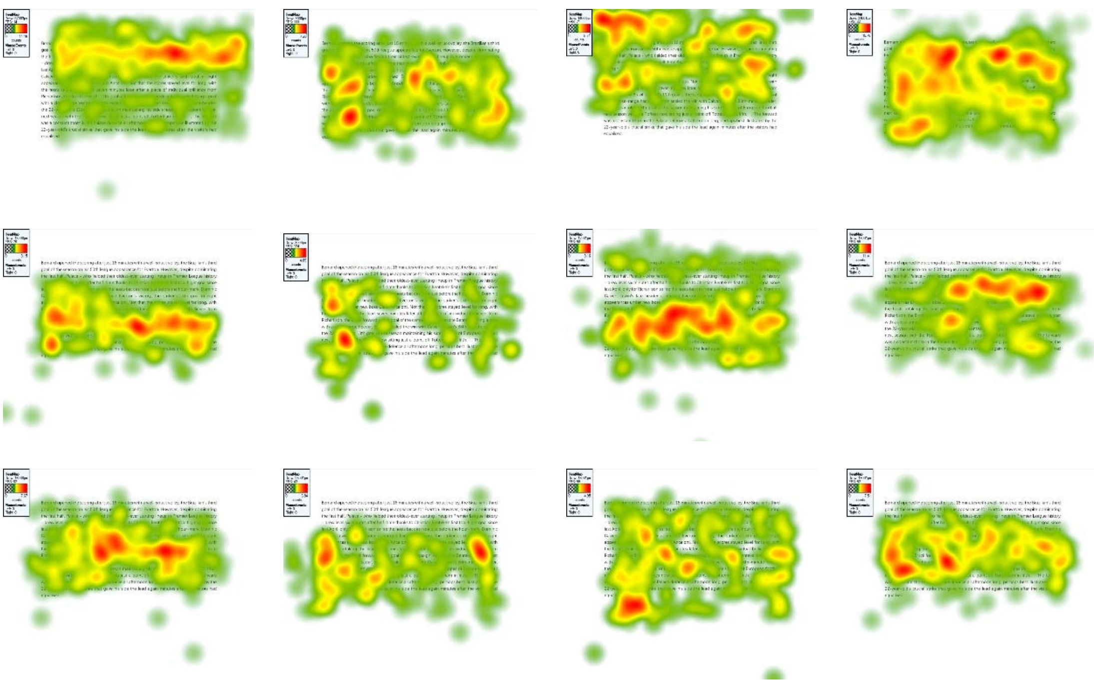


Figure 80: Heat Map for summary showing pattern of fixations for 12 participants

As figure 81 shows, participants began their reading with a longer fixation in the image of the first panel. We divided the panels into two sections: the upper part containing the team names, time and score; the lower part containing the pitch, players, and scorer's name. Second, the time spent on the first panel was higher than on the other panels, with a decreasing time from the top to the bottom panels. While the Z pattern makes sense in a narrative comic, where the flow of the story could encourage a left-to-right, top-to-bottom scan path, our comic does not contain a ‘narrative’ so much as a temporal sequence of events. As such, the scan path might be influenced by making sense of the soccer game.

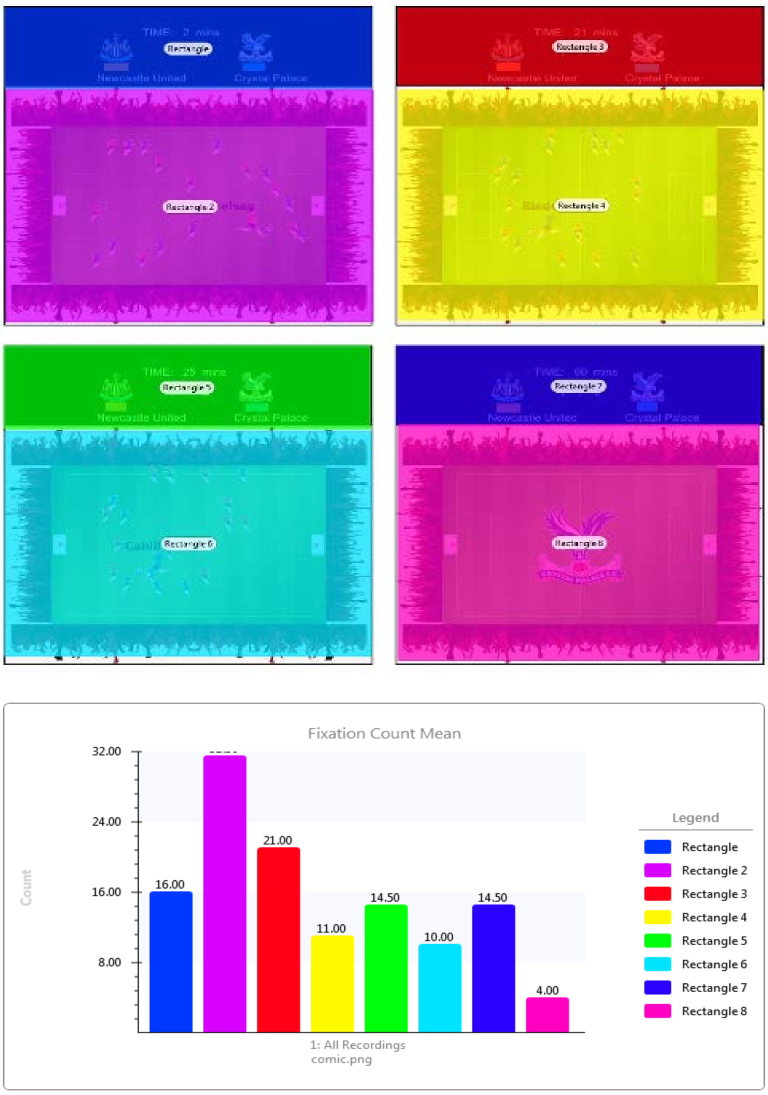
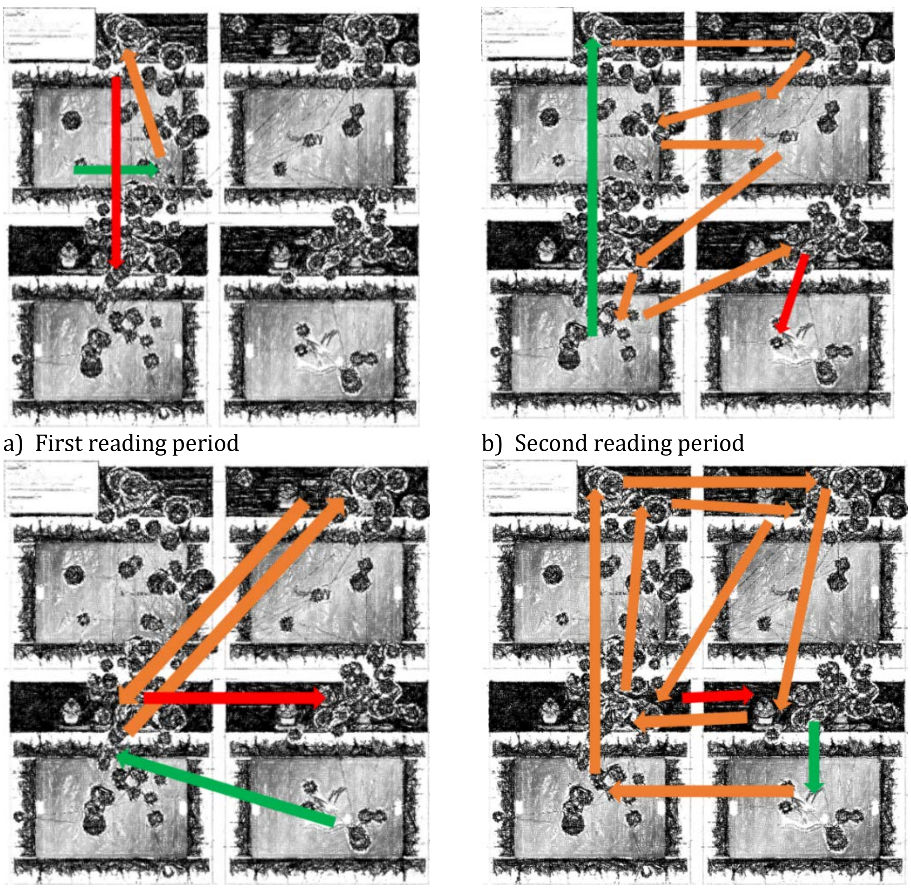


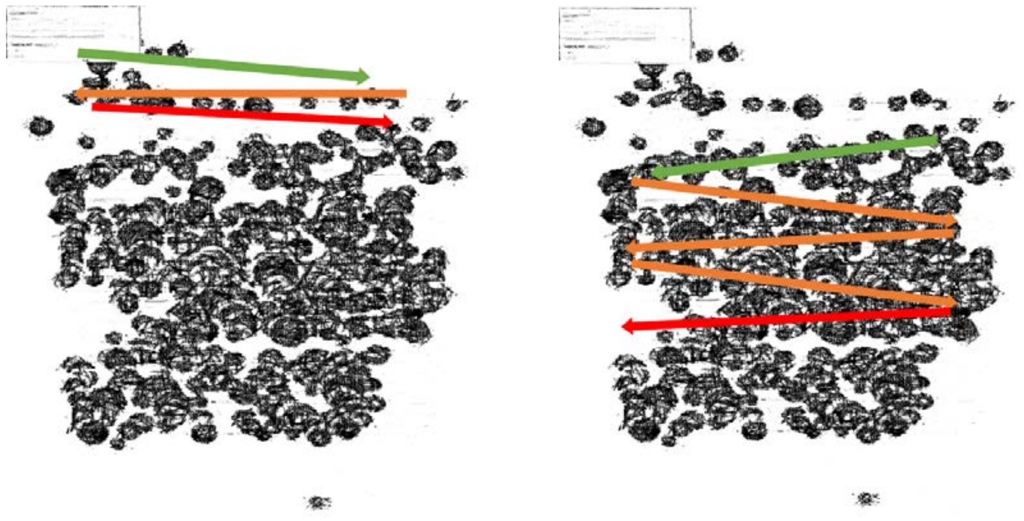
Figure 81: The Fixation data for Upper and Lower parts of each panel

In figure 82, we divided the reading activity into four periods. These roughly corresponded to the viewing of each panel. It is clear that gaze transition is between panels. In this example, the participant began in the center of the image of panel 1, and then moved directly down to panel 3. In the second period b, we found that when the reader moved back to the first panel, fixating on the upper part of panel 1 to the score. From this, the scan path moves left-to-right and diagonally down across panels 2 and 3 to panel 4 (in a sequence that approximates a 'Z'). This path was reinforced in the third period, we observed the participants retraced the path back to panel 2. This might reflect some uncertainty about the previously scanned information, requiring a re-reading of panels 2 and 3. In the fourth period, the participant scanned all the panels in a loop, which implied a final refresh of the information. In this respect, reading the comic was less about deriving a story than acquiring information.



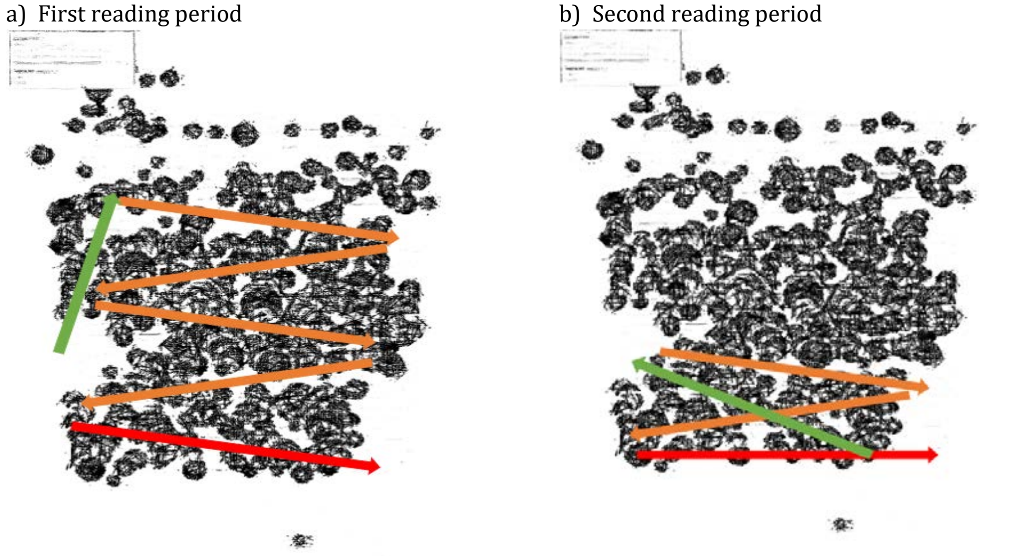
a) First reading period
b) Second reading period
c) Third reading period
d) Final reading period

Figure 82: Reading patterns for Comic(green arrow indicates initial scan transition; red indicates final scan transition; amber indicates intermediate scan paths).



a) First reading period

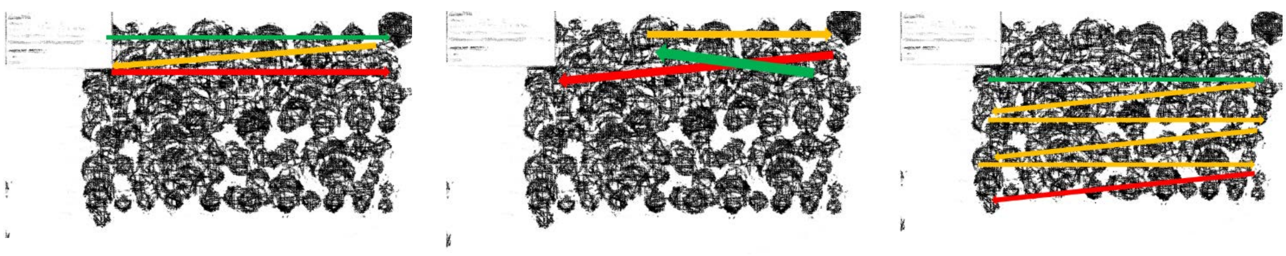
b) Second reading period



c) Third reading period

d) Final reading period

Figure 83: Reading patterns for full text (green arrow indicates initial scan transition; red indicates final scan transition; amber indicates intermediate scan paths).



a) First reading period

b) Second reading period

c) Final reading period

Figure 84: Reading patterns for summary (green arrow indicates initial scan transition; red indicates final scan transition; amber indicates intermediate scan paths).

In Figures 83 and 84, we performed a phased analysis of the reading trajectories of the text. It was clear that gaze transitions were between the contents of the text. In this example, the participant's first sight was the beginning of the text. Then the participant's gaze moved down continuously, and the trajectories of the gaze points were in a sequence that followed a 'Z'. This means that scans for textual information were in accordance with the 'Z' scans mentioned in the introduction. At the same time, the participants would backtrack the information that had been scanned, and the backtracking means that the participants obtained key information from it and focus on it, which corresponded to the hot spots in Figure 78 and Figure 79.

Through analysis of gaze trajectories, we found that for text gaze, participants still followed the ‘Z’ saccade. But for the gaze trajectories of comics, we found that the gaze points are more focused on the locations that contain key information. While the gaze point was jumpy in the comic, the trajectory approximately conformed to a ‘Z’ scan. This is consistent with the assumptions we mentioned in the introductory section.

5.3 Conclusions for Eye-Tracking Experiment

Surprisingly, very few participants answered all questions correctly, despite there being only 7 questions each with a choice of 3 possible answers. We had informed participants that we would ask them questions after their reading (and that the purpose of reading the material would be to understand it sufficiently to answer questions). However, the full-text and summary text contained material that the questions did not include. As such, this material might have seemed distracting or redundant to the task at hand. Reading to filter the text without a clear notion of what defines 'salience' (i.e., relevant to answering the questions) proved challenging. We might have expected participants who read the comic first to have an idea of what information to look for in the text, and thus, we might have expected an order effect in which the third trial would yield a higher score than the first trial. This was not necessarily the case, and the results seemed more influenced by the medium than the trial.

Eye movement is task-dependent (Mikkonen, 2019; Ballard, 2009; Cohn, 2019). Consequently, we expected participants who read the comic to attend to specific details. Interestingly, the scan path followed the well-known Z-path. This suggests that their reading of the comic was as much influenced by the layout (in a grid) as by the task (of seeking information). This also means that participants were not merely attending to the textual content to answer the questions but also looking at the images. Of course, embedding text in the image (i.e., the name of the goal scorer) would contribute to this pattern. As we can see (from figure 80), fixations tend toward the text. However, from figure 80, we also observe that certain elements of the images attract attention, and the scan-paths (figure 82) suggest that participants are looking at the comic as a combination of graphics and text. At the end of the trial, when we asked participants which medium they preferred, they uniformly agreed that the comic was easier and more enjoyable to read.

In conversations with participants after the experiment, it became apparent that the visual appearance of the comic was more appealing than the blocks of text. While the graphic design of our comics is rudimentary and we could improve it, the fact that we could construct these comics using simple layout principles and a limited image dictionary suggests that we can extend this approach. Moreover, the rules that the story grammar captures seem to enable a sufficient range of events for soccer.

6 Conclusion and Discussion

As mentioned earlier, most text summarization relies on statistical analysis methods, and most visualizations are data visualizations. Text visualization exhibits significant limitations within the realm of deep learning. We propose a unique method for converting text into comics based on story grammar, focusing on visualizing sports coverage. By defining story grammar, we generate domain-specific lexicons that provide robust data support for our comic system. We utilize Spacy's fast natural language analysis function to quickly summarize and analyze the text that our crawler system extracts, which significantly improves the system's operational efficiency. We also take advantage of the correlation between story grammar and vocabulary to visualize events. This system demonstrates the close relationship between language grammar and story grammar. The visualization of text data offers fresh perspectives for text summarization and visualization of complex texts. Our text visualization method differs from previous methods that only visualized data, as well as from the complex structures of deep learning and extensive manual labeling work. We build a lexicon through story grammar, which not only reduces human costs but also enhances the accuracy and coherence of text summarization. Simultaneously, our method can visualize and incorporate multiple events within a text into a comprehensive comic. This approach also avoids the limitations of inaccurate visual image expression and the inability to segment and visualize the entire text found in deep learning models. Users can understand the important information within the text through images, eliminating the need to spend too much time reading the original text. Our case study's results demonstrate the effectiveness of our proposed story grammar and caricature generation, further highlighting its potential for future research and development.

6.1 Research Question 1: What is the most appropriate way to summarise text to render information in the form of comics?

Using Story Grammar assists in text abstraction. It minimizes the need for complex Machine Learning while maintaining consistent interpretation for the domain. Specifically, the story grammar outlines a framework for interpreting salient events in the domain of a specific sport, aiding dictionary definition. In any sports report, numerous synonyms for common events exist (likely to reduce repetition). This could mean, for instance, that a 'goal' scored in soccer might not be explicitly mentioned in the text; the report could say a team 'drew level score 2 minutes later.' This requires understanding that (a) 'drew level' indicates a score change, (b) the timing of the event relative to a previous event, and (c) identifying the player who scored. While Machine Learning could discover each of these, combining story grammar (to constrain the search elements) and dictionary (to limit synonyms) simplifies the task.

The application of story grammar also provides robust structural support for the visualization system. Through the story grammar and its defined special dictionary, the system can more quickly and accurately retrieve required image elements from the image library and combine them into meaningful pictures. The story grammar and dictionary assist the system in improving event vocabulary recognition accuracy, thereby reducing interference factors caused by polysemous words and enhancing text visualization accuracy. This means that story grammar can help the system significantly increase operational speed while reducing the error rate of generation.

6.2 Research Question 2 & 3: Is there any difference between the original text, abstract text and comic material for readers? Which of the three materials, the original text, the summary text, and the comic material, can help readers understand the events described?

Two experiments in this study investigated the effects of different forms of information delivery (raw text, abstracts, cartoons) on reading comprehension, memory, and user experience. The results in the web experiment showed that although the reading time of the original material was longer, participants had better comprehension and retention when reading the comic, and were more likely to recommend it to a friend. In addition, although the reading time of comics is similar to that of summaries, they perform better in terms of comprehension, memory, and user experience.

Eye-tracking experiments explore reading strategies for text and comics. The study found that while participants may have been confused and distracted when reading the original text and abstract, they were better able to focus on important information when reading the cartoon, which may be related to the cartoon's visual representation. The comics help participants develop a comprehensive understanding of the material through visual images, thereby enhancing their comprehension and retention.

In terms of how well they understood the information, participants did better on literal comprehension than on evaluative or inferential comprehension. This may be because literal comprehension requires only the identification and recall of important material, whereas evaluative and inferential comprehension requires in-depth analysis and comprehension of the material. The study also found that when reading the comic, participants excelled not only in literal comprehension, but also in assessment and inferential comprehension, possibly due to the fact that the comic's visual representation facilitated participants' comprehension and memory.

In terms of user experience, participants prefer to read comic, think comic is easier to read, and are more willing to recommend it to friends. Although participants rated the summary's ease of reading closer to that of the comic, they rated the summary as worse at presenting important information, suggesting that they were aware of their comprehension and memory abilities while reading.

This research provides us with valuable insights into how information delivery affects reading comprehension, memory, and user experience, and provides some directions for how information delivery can be optimized in the future, for example, using the visual representation of comics to increase the attractiveness of information power and understanding.

In conclusion, the comic material's readability is the highest, followed by the summary material, and finally the original text. Therefore, our prior assumption that comics can reduce redundant events and improve comprehension efficiency seems feasible.

6.3 Problems and challenges arising from this research

While designing this subject, we combined story grammar and natural language grammar to extract text summaries. Since we only used story grammar, we cannot conclude with certainty that our method is the "best" way to summarise the text. Spacy, as the de facto standard in NLP, provides reliable source material for analysis. Using NLP theory to extract text ensures the greatest coherence of text information and sentence readability. However, this method's limitation is that we need to manually analyse the registration of different types of events to manually extract the necessary dictionary and input it into the system. If we create a dictionary for other events (like basketball games), we will need to recreate the new dictionary content. Therefore, we need to further improve the dictionary creation part. We believe that in the process of building the dictionary, we could use machine learning based on natural language recognition theory to extract the words in the text and classify them into a dictionary according to the story grammar structure. Using machine learning functions will help the system automate the establishment of dictionaries, thereby saving human resources. Simultaneously, we can apply machine learning to more different types of check-in so that we can achieve the goal of automatically establishing a dictionary by adjusting the relevant parameters.

We have completed the project's system prototype, and it can adapt to the vast majority of football matches (with occasional minor adjustments). We have also preliminarily completed the design of the picture library, but the current picture library cannot change the colour according to the different teams. Therefore, we need to add image libraries and codes that can change colour based on team information to the system. But the current design is sufficient for the current experimental investigation.

In the survey experiment part, the number of participants in the network experiment was limited. Still, each material set was repeated twice by participants who didn't understand, and the current experimental results align with the target hypothesis. However, the number of repetitions is far from enough. Therefore, we should increase the number of participants in the network experiment to increase the number of repetitions of each material group to obtain more accurate experimental results. Finally, we can verify that the project system can achieve the expected goals. For the eye-tracking experiment, while the number of 12 participants is acceptable, we still need more participants to conduct the experiment, so that we can use more experimental data to obtain accurate experimental results. Simultaneously, for the two experiments, we can verify individual differences by asking them whether they have the basic theoretical knowledge of this type of sports, reading habits, and answering strategies.

6.4 Future developments

We mentioned in our discussion of story grammars that sports with a more significant 'tactical' component might not be well suited to the event-based approach we have taken in this thesis. Let's speculate on how we might incorporate tactics into the story grammar and propose the following additions and extensions. We could include panels to show the team line-up at critical points in the game, similar to how coaches visualise game plans for specific moves in a game. For a free-flowing sport like soccer, placing the teams in 'realistic' positions when a goal is scored might suffice. If we know, for example, that a team favours a specific formation, like 4-4-2 or 3-4-3, we could use this to arrange the players in a panel. If we believe that the attacking team has pulled the defending team to one side of the pitch, leaving space for an attack on the opposite side, we could indicate this. Sometimes, the original report will provide enough information, but in other cases, we would have to make assumptions about the teams and players' behaviour.

We might also include narrative elements in our comics. For instance, comic stories often use 'flashbacks' to provide context or motivation for a character's mood or actions. As we noted in our summary (figure 36), newspaper reports include historical information, such as the last time the two teams played, or contextual information (like whether one team had a run of losses prior to the match and if this meant the manager was under pressure). Showing previous matches or a worried-looking manager in a panel could provide such context.

The possibility of adding panels to the comic to provide contextual information is related to the choice of how many 'events' to include. In other words, one might wish to view the panel on a small screen (Augereau, 2016), in which case the number of panels could be constrained to a single panel at a time or to the smallest number of panels that would allow text to be read. In this case, the reader might have a notion of how many panels would be acceptable. We suspect that this number could depend not only on-screen size but also on the purpose of viewing the comic. Someone wishing to enjoy a detailed match report in comic form might accept many panels with lots of historical and contextual detail

Linking our constructed comics to the original text allows us to question what information the transition loses. The text phrases like "tight angle" and "crashed in off the post" are not represented in the current comic version. However, we could accommodate these with more images in our image library. For longer passages, there might be information loss, and we are exploring metrics to define this, such as based on information theory. However, some information might not relate directly to the match events and could provide background information about a team's recent performance or its position in the league table. An equally interesting question is what the images might 'add' to the match account that the text omits. While we are unlikely to represent elements not mentioned in the text, providing a spectators background behind the goal in our panels could offer a sense of a 'live' match. And one could imagine that the image library could present different football stadiums to enhance this.

Our focus was on visualising sports reports, but it's plausible to assume that we could extend this approach to other information conforming to a story grammar, i.e., a sequence of events where characters seek to achieve end-goals and respond to situational demands. Furthermore, if we can define a generic story grammar for a specific domain, we could also use the comics to generate 'alternative' what-if versions of the story. While sports fans might enjoy a story where their team wins the trophy instead of coming off second best, a more useful version of this could be in terms of Intelligence Analysis. Here, a comic could visualise timelines and dependencies to support Analysis of Competing Hypotheses (Pherson and Heuer Jr, 2020) through the creation of branching narratives (Eisner, 2008a) generated by modifying elements of the story grammar.

Reference

Albrengues, C., Lavigne, F., Aguilar, C., Castet, E. and Vitu, F. (2019) ‘Linguistic processes do not beat visuo-motor constraints, but they modulate where the eyes move regardless of word boundaries: Evidence against top-down word-based eye-movement control during reading’, *PLoS One*, 14(7), p. e0219666.

Alves, T., McMichael, A., Simões, A., Vala, M., Paiva, A. and Aylett, R. (2007) *Comics2D: Describing and creating comics from story-based applications with autonomous characters*. Available at: https://www.researchgate.net/profile/Ruth-Aylett/publication/228359599_Comics2D_Describing_and_creating_comics_from_story-based_applications_with_autonomous_characters/links/0fcfd50697c84d3dcc000000/Comics2D-Describing-and-creating-comics-from-story-based-applications-with-autonomous-characters.pdf (Accessed: 18 July 2023).

Augereau, O., Matsubara, M. and Kise, K. (2016) ‘Comic visualization on smartphones based on eye tracking’, *Proceedings of the 1st International Workshop on coMics ANalysis, Processing and Understanding*. Cancun, Mexico, December 04. Association for Computing Machinery.

Bach, B., Riche, N. H., Carpendale, S. and Pfister, H. (2017) ‘The Emerging Genre of Data Comics’, *IEEE Computer Graphics and Applications*, 37(3), pp. 6-13.

Bach, B., Wang, Z., Farinella, M., Murray-Rust, D. and Riche, N. H. (2018) ‘Design Patterns for Data Comics’, *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. Montreal QC, Canada, April 4-9, 2009. Association for Computing Machinery.

Ballard, D. H. (2009) ‘Modeling the role of task in the control of gaze’, *Visual Cognition*, 17(6-7), pp. 1185-1204.

Bartlett, F. A. and Remembering, A. (1932) *A study in experimental and social psychology*. New York: Cambridge University Press.

Basaraba, D., Yovanoff, P., Alonzo, J. and Tindal, G. (2013) ‘Examining the structure of reading comprehension: do literal, inferential, and evaluative comprehension truly exist?’, *Reading & Writing*, 26(3), pp. 349-379.

Bateman, S., Mandryk, R. L., Gutwin, C., Genest, A., McDine, D. and Brooks, C. (2010) ‘Useful junk?

the effects of visual embellishment on comprehension and memorability of charts', *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Atlanta, Georgia, USA, April 10-15, 2010. Association for Computing Machinery.

Bengio, Y., Mesnil, G., Dauphin, Y. and Rifai, S. (2013) 'Better mixing via deep representations', *Proceedings of the 30th International Conference on International Conference on Machine Learning-Volume 28*. Atlanta, GA, USA. JMLR.org.

Bereiter, C. (1983) 'Story grammar as knowledge', *Behavioral & Brain Sciences*, 6(4), pp. 593-594.

Berger, A. and Mittal, V. O. (2000) 'Query-relevant summarization using FAQs', *Proceedings of the 38th Annual Meeting on Association for Computational Linguistics*. Hong Kong. Association for Computational Linguistics.

Black, J. B. and Wilensky, R. (1979) 'An evaluation of story grammars', *Cognitive Science*, 3(3), pp. 213-229.

Booker, C. (2004) *The seven basic plots: Why we tell stories*. London: A&C Black.

Borkin, M. A., Vo, A. A., Bylinskii, Z., Isola, P., Sunkavalli, S., Oliva, A. and Pfister, H. (2013) 'What makes a visualization memorable?', *IEEE Transactions on Visualization and Computer Graphics*, 19(12), pp. 2306-2315.

Brewer, W. F. and Lichtenstein, E. H. (1980) *Event schemas, story schemas, and story grammars*. Champaign, Ill. [Online]. Available at: https://cir.nii.ac.jp/crid/1571417125580552064.

Brooks, P. H. (1977) 'The role of action lines in children's memory for pictures', *Journal of Experimental Child Psychology*, 23(1), pp. 98-107.

Brown, P. F., Della Pietra, V. J., Desouza, P. V., Lai, J. C. and Mercer, R. L. (1992) 'Class-based n-gram models of natural language', *Computational Linguistics*, 18(4), pp. 467-480.

Burr, D. C. and Ross, J. (2002) 'Direct evidence that "speedlines" influence motion mechanisms', *Journal of Neuroscience*, 22(19), pp. 8661-8664.

Buscher, G., Cutrell, E. and Morris, M. R. (2009) 'What do you see when you're surfing? using eye tracking to predict salient regions of web pages', *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Boston, MA, USA. Association for Computing Machinery.

Carbonell, J. and Goldstein, J. (2017) 'The Use of MMR, Diversity-Based Reranking for Reordering Documents and Producing Summaries', *SIGIR Forum*, 51(2), pp. 209–210.

Chowdhury, G. G. (2003) ‘Natural language processing’, *Annual Review of Information Science and Technology*, 37, pp. 51-89.

Carroll, J. B. (1964) ‘Language and thought’. Englewood Cliffs, NJ: Prentice-Hall.

Cohn, N. (2010) ‘Extra! Extra! Semantics in comics!: The conceptual structure of Chicago Tribune advertisements’, *Journal of Pragmatics*, 42(11), pp. 3138-3146.

Cohn, N. (2013a) ‘Navigating comics: An empirical and theoretical approach to strategies of reading comic page layouts’, *Frontiers in Psychology*, 4, p. 186.

Cohn, N. (2013b) *The Visual Language of Comics: Introduction to the Structure and Cognition of Sequential Images*. London: A&C Black.

Cohn, N. (2013c) ‘Visual narrative structure’, *Cognitive Science*, 37(3), pp. 413-452.

Cohn, N. (2015) ‘Narrative conjunction's junction function: The interface of narrative grammar and semantics in sequential images’, *Journal of Pragmatics*, 88, pp. 105-132.

Cohn, N. (2019) ‘Your Brain on Comics: A Cognitive Model of Visual Narrative Comprehension’, *Topics in Cognitive Science*, 12(1), pp. 352-386.

Cohn, N. (2020) ‘Your brain on comics: A cognitive model of visual narrative comprehension’, *Topics in Cognitive Science*, 12(1), pp. 352-386.

Cowan, N. (2008) ‘What are the differences between long-term, short-term, and working memory?’, *Progress in Brain Research*, 169, pp. 323-338.

Cui, W., Zhang, X., Wang, Y., Huang, H., Chen, B., Fang, L., Zhang, H., Lou, J. G. and Zhang, D. (2020) ‘Text-to-Viz: Automatic Generation of Infographics from Proportion-Related Natural Language Statements’, *IEEE Transactions on Visualization and Computer Graphics*, 26(1), pp. 906-916.

Czerwiec, M. K. (2018) ‘Representing AIDS in comics’, *AMA Journal of Ethics*, 20(2), pp. 199-205.

Djamasbi, S., Siegel, M. and Tullis, T. (2011) ‘Visual Hierarchy and Viewing Behavior: An Eye Tracking Study’, *International Conference on Human-Computer Interaction*. Berlin, Heidelberg. Springer.

Eisner, W. (2008a) *Comics and sequential art: Principles and practices from the legendary cartoonist*. New York: WW Norton & Company.

Eisner, W. (2008b) *Graphic storytelling and visual narrative*. New York: WW Norton & Company.

Epelboim, J., Booth, J. R. and Steinman, R. M. (1994) ‘Reading unspaced text: Implications for theories of reading eye movements’, *Vision Research*, 34(13), pp. 1735-1766.

Fattah, M. A. and Ren, F. (2009) ‘GA, MR, FFNN, PNN and GMM based models for automatic text summarization’, *Computer Speech & Language*, 23(1), pp. 126-144.

Fink, A. (2003) *How to design survey studies*. Thousand Oaks, CA: Sage.

Foulsham, T., Wybrow, D. and Cohn, N. (2016) ‘Reading without words: Eye movements in the comprehension of comic strips’, *Applied Cognitive Psychology*, 30(4), pp. 566-579.

Franks, I. M. (1986) ‘Eyewitness testimony in sport’, *Journal of Sport Behavior*, 9(1), pp. 38-45.

Gambhir, M. and Gupta, V. (2017) ‘Recent automatic text summarization techniques: a survey’, *Artificial Intelligence Review*, 47, pp. 1-66.

Goldsberry, K. (2012) ‘Courtvision: New visual and spatial analytics for the nba’, *Proceedings of 6th Annual MIT Sloan Sports Analytics Conference*. http://www.sloansportsconference.com/wp-content/uploads/2012/02/Goldsberry_Sloan_Submission.pdf

Goodman, K. S. (1976) 'Behind the eye: What happens in reading', in Singer, H. and Ruddell, R. B. (eds.) *Theoretical models and processes of reading*. Newark, DE: International Reading Association, pp. 470-496.

Gough, P. B. (1972) ‘One Second of Reading’, *Visible Language*, 6(4), pp. 291-320.

Gudmundsson, J. and Horton, M. (2017) ‘Spatio-Temporal Analysis of Team Sports’, *ACM Computing Surveys*, 50(2), pp. 1–34.

Haroz, S., Kosara, R. and Franconeri, S. L. (2015) ‘ISOTYPE Visualization: Working Memory, Performance, and Engagement with Pictographs’, *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. Seoul, Republic of Korea, April 18-23, 2015. Association for Computing Machinery.

Harrison, L., Reinecke, K. and Chang, R. (2015) ‘Infographic Aesthetics: Designing for the First Impression’, *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. Seoul, Republic of Korea, April 18-23, 2015. Association for Computing Machinery.

Hegarty, M. and Just, M.-A. (1993) ‘Constructing mental models of machines from text and diagrams’, *Journal of Memory and Language*, 32(6), pp. 717-742.

Holsanova, J., Holmberg, N. and Holmqvist, K. (2009) ‘Reading information graphics: The role of spatial contiguity and dual attentional guidance’, *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 23(9), pp. 1215-1226.

Holsanova, J., Rahm, H. and Holmqvist, K. (2006) ‘Entry points and reading paths on newspaper spreads: comparing a semiotic analysis with eye-tracking measurements’, *Visual Communication*, 5(1), pp. 65-93.

Hoover, W. A. and Tunmer, W. E. (1993) 'The components of reading', in Thompson, G. B., Tunmer, W. E. and Nicholson, T. (eds.) *Reading acquisition processes*. Clevedon, UK: Multilingual Matters, pp. 1-19.

Hosler, J. and Boomer, K. B. (2011) ‘Are comic books an effective way to engage nonmajors in learning and appreciating science?’, *CBE—Life Sciences Education*, 10(3), pp. 309-317.

Hudson, T. (1998) ‘Theoretical perspectives on reading’, *Annual Review of Applied Linguistics*, 18, pp. 43-60.

Hullman, J., Adar, E. and Shah, P. (2011) ‘The impact of social information on visual judgments’, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Vancouver, BC, Canada, May 7-12, 2011. Association for Computing Machinery.

Jeong, H. (2012) ‘A comparison of the influence of electronic books and paper books on reading comprehension, eye fatigue, and perception’, *The Electronic Library*, 30(3), pp. 390-408.

Johnson-Laird, P. N., Byrne, R. M. and Schaeken, W. (1992) ‘Propositional reasoning by model’, *Psychological Review*, 99(3), p. 418.

Kausar, M. A., Dhaka, V. S. and Singh, S. K. (2013) ‘Web crawler: a review’, *International Journal of Computer Applications*, 63(2), pp. 31-36.

Kawabe, T. and Miura, K. (2006) ‘Representation of dynamic events triggered by motion lines and static human postures’, *Experimental Brain Research*, 175(2), pp. 372-375.

Kim, H. and Francis, G. (1998) ‘A computational and perceptual account of motion lines’, *Perception*, 27(7), pp. 785-797.

Kim, N. W., Schweickart, E., Liu, Z., Dontcheva, M., Li, W., Popovic, J. and Pfister, H. (2016) ‘Data-driven guides: Supporting expressive design for information graphics’, *IEEE Transactions on Visualization and Computer Graphics*, 23(1), pp. 491-500.

Kucher, K. and Kerren, A. (2015) ‘Text visualization techniques: Taxonomy, visual survey, and community insights', *2015 IEEE Pacific Visualization Symposium (PacificVis)*. Hangzhou, China, 14-17 April 2015. IEEE.

Kurzhals, K., Fisher, B., Burch, M. and Weiskopf, D. (2016) 'Eye tracking evaluation of visual analytics', *Information Visualization*, 15(4), pp. 340-358.

Levie, W. H. and Lentz, R. (1982) 'Effects of text illustrations: A review of research', *ECTJ*, 30(4), pp. 195-232.

Levin, J. R. (1987) 'On empirically validating functions of pictures in prose', *The Psychology of Illustration*, 1, pp. 51-86.

Lin, S.-F., Lin, H.-s., Lee, L. and Yore, L. D. (2015) 'Are science comics a good medium for science communication? The case for public learning of nanotechnology', *International Journal of Science Education, Part B*, 5(3), pp. 276-294.

Liu, J. (2004) 'Effects of comic strips on L2 learners' reading comprehension', *TESOL Quarterly*, 38(2), pp. 225-243.

Liu, S., Zhou, M. X., Pan, S., Qian, W., Cai, W. and Lian, X. (2009) 'Interactive, topic-based visual text summarization and analysis', *Proceedings of the 18th ACM conference on Information and knowledge management*. Hong Kong, China, November 2-6, 2009. Association for Computing Machinery.

Loper, E. (2004) *NLTK: Building a pedagogical toolkit in Python*. Available at: https://www.semanticscholar.org/paper/NLTK%3A-Building-a-Pedagogical-Toolkit-in-Python-Loper/5d133dad3ac0d392f6f176a5ea457f614dd2538a?p2df (Accessed: 19 July 2023).

Loper, E. and Bird, S. (2002) 'NLTK: the Natural Language Toolkit', *Proceedings of the ACL-02 Workshop on Effective tools and methodologies for teaching natural language processing and computational linguistics-Volume 1*. Philadelphia, Pennsylvania, 7 July 2002. Association for Computational Linguistics.

Ma, K. L., Liao, I., Frazier, J., Hauser, H. and Kostis, H. N. (2012) 'Scientific Storytelling Using Visualization', *IEEE Computer Graphics & Applications*, 32(1), pp. 12-19.

Mandler, J. M. and Johnson, N. S. (1977) 'Remembrance of things parsed: Story structure and recall', *Cognitive Psychology*, 9(1), pp. 111-151.

McCloud, E. S. (1993) *The Invisible Art*. New York: Harper Collins.

McDermott, J. E., Partridge, M. and Bromberg, Y. (2018) 'Ten simple rules for drawing scientific comics ', *PLoS Computational Biology*, 14(1), p. e1005845.

Meuschke, M., Garrison, L. A., Smit, N. N., Bach, B., Mittenentzwei, S., Weiß, V., Bruckner, S., Lawonn, K. and Preim, B. (2022) 'Narrative medical visualization to communicate disease data', *Computers & Graphics*, 107, pp. 144-157.

Mikkonen, K. and Lautenbacher, O. P. (2019) 'Global Attention in Reading Comics: Eye movement indications of interplay between narrative content and layout', *ImageTexT*, 10(2).

Mikolov, T., Chen, K., Corrado, G. and Dean, J. (2013) *Efficient Estimation of Word Representations in Vector Space*. Available at: https://arxiv.org/abs/1301.3781 (Accessed: 19 July 2023).

Miller, G. A. (1995) 'WordNet: a lexical database for English', *Communications of the ACM*, 38(11), pp. 39-41.

Nomoto, T. and Matsumoto, Y. (2001) 'A new approach to unsupervised text summarization', *Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval*. New Orleans, Louisiana, USA, 01 September 2001. Association for Computing Machinery.

Perfetti, C. A. (1995) 'Cognitive Research Can Inform Reading Education', *Journal of Research in Reading*, 18(2), pp. 106-15.

Perin, C., Vuillemot, R. and Fekete, J. D. (2013) 'SoccerStories: A Kick-off for Visual Soccer Analysis', *IEEE Transactions on Visualization and Computer Graphics*, 19(12), pp. 2506-2515.

Pherson, R. H. and Heuer Jr, R. J. (2020) *Structured analytic techniques for intelligence analysis*. Washington, DC: CQ Press.

Pomorska, K. (1985) 'Poetics of prose', *Verbal art, verbal sign, verbal time*, pp. 65-75.

Postal, P. (1964) 'Underlying and Superficial Linguistic Structure', *Harvard Educational Review*, 34(2), pp. 246-266.

Ranker, J. (2007) 'Using comic books as read‐alouds: Insights on reading instruction from an English as a second language classroom', *The Reading Teacher*, 61(4), pp. 296-305.

Ratinov, L. and Roth, D. (2009) 'Design challenges and misconceptions in named entity recognition', *Proceedings of the Thirteenth Conference on Computational Natural Language Learning*. Boulder, Colorado, June 4-5, 2009. Association for Computational Linguistics.

Rayner, K., Rotello, C. M., Stewart, A. J., Keir, J. and Duffy, S. A. (2001) ‘Integrating Text and Pictorial Information: Eye Movements When Looking at Print Advertisements’, *Journal of Experimental Psychology Applied*, 7(3), pp. 219-226.

Reed, S., Akata, Z., Yan, X., Logeswaran, L., Schiele, B. and Lee, H. (2016) ‘Generative adversarial text to image synthesis’, *Proceedings of the 33rd International Conference on International Conference on Machine Learning - Volume 48*. New York, NY, USA, June 19-24, 2016. JMLR.org.

Reep, C. and Benjamin, B. (1968) ‘Skill and chance in association football’, *Journal of the Royal Statistical Society. Series A (General)*, 131(4), pp. 581-585.

Rumelhart, D. E. (1975) 'Notes on a schema for stories', in Bobrow, D. G. and Collins, A. (eds.) *Representation and Understanding*. San Diego: Morgan Kaufmann, pp. 211-236.

Satyanarayan, A. and Heer, J. (2014) ‘Lyra: an interactive visualization design environment’, *Proceedings of the 16th Eurographics Conference on Visualization*. Swansea, Wales, United Kingdom, June 9-13, 2014. Eurographics Association.

Scharlach, T. D. (2011) ‘START Comprehending: Students and Teachers Actively Reading Text’, *The Reading Teacher*, 62(1), pp. 20-31.

Segel, E. and Heer, J. (2010) ‘Narrative visualization: Telling stories with data’, *IEEE Transactions on Visualization and Computer Graphics*, 16(6), pp. 1139-1148.

Skau, D. and Kosara, R. (2017) ‘Readability and precision in pictorial bar charts’, *Proceedings of the Eurographics/IEEE VGTC Conference on Visualization: Short Papers*. Barcelona, Spain, June 12-16, 2017. Eurographics Association.

Smith, F. (2012) *Understanding reading: A psycholinguistic analysis of reading and learning to read*. New York: Routledge.

Stein, M., Janetzko, H., Lamprecht, A., Breitkreutz, T., Zimmermann, P., Goldlücke, B., Schreck, T., Andrienko, G., Grossniklaus, M. and Keim, D. A. (2018) ‘Bring It to the Pitch: Combining Video and Movement Data to Enhance Team Sport Analysis’, *IEEE Transactions on Visualization and Computer Graphics*, 24(1), pp. 13-22.

Strzalkowski, T., Wang, J. and Wise, B. (1998) ‘A robust practical text summarization’, *Proceedings of the AAAI Symposium on Intelligent Text Summarization*. 1998. AAAI Press.

Tas, O. and Kiyani, F. (2007) ‘A survey automatic text summarization’, *PressAcademia Procedia*, 5(1), pp. 205-213.

Thompson, G. B., Tunmer, W. E. and Nicholson, T. (1993) *Reading acquisition processes*. Clevedon, UK: Multilingual Matters.

Thorndyke, P. W. (1977) ‘Cognitive structures in comprehension and memory of narrative discourse’, *Cognitive Psychology*, 9(1), pp. 77-110.

Tsujimoto, S. and Asada, H. (1992) ‘Major components of a complete text reading system’, *Proceedings of the IEEE*, 80(7), pp. 1133-1149.

Tversky, B., Morrison, J. B. and Betrancourt, M. (2002) ‘Animation: Does it facilitate learning’, *International Journal of Human-Computer Studies*, 57, pp. 247-262.

Van De Rakt, M. (2022) *The flesch reading ease score: why and how to use it*. Available at: https://yoast.com/flesch-reading-ease-score/ (Accessed: 30 May 2023).

Van Rossum, G. and Drake, F. L. (2003) *An introduction to Python*. Bristol: Network Theory Ltd.

VanRossum, G. and Drake, F. L. (2010) *The python language reference*. Amsterdam, Netherlands: Python Software Foundation.

Versaci, R. (2001) ‘How comic books can change the way our students see literature: One teacher's perspective’, *The English Journal*, 91(2), pp. 61-67.

Viégas, F. B., Golder, S. and Donath, J. (2006) ‘Visualizing email content: portraying relationships from conversational histories’, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Montréal, Québec, Canada, April 22-27, 2006. Association for Computing Machinery.

Von der Malsburg, T., Kliegl, R. and Vasishth, S. (2015) ‘Determinants of scanpath regularity in reading’, *Cognitive Science*, 39(7), pp. 1675-1703.

Wang, Y., Zhang, H., Huang, H., Chen, X., Yin, Q., Hou, Z., Zhang, D., Luo, Q. and Qu, H. (2018) ‘InfoNice: Easy Creation of Information Graphics’, *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. Montreal QC, Canada, April 21-26, 2018. Association for Computing Machinery.

Wang, Z., Dingwall, H. and Bach, B. (2019) ‘Teaching Data Visualization and Storytelling with Data Comic Workshops’, *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. Glasgow, Scotland UK, May 4-9, 2019. Association for Computing

Machinery.

Wang, Z., Ritchie, J., Zhou, J., Chevalier, F. and Bach, B. (2020) ‘Data comics for reporting controlled user studies in human-computer interaction’, *IEEE Transactions on Visualization and Computer Graphics*, 27(2), pp. 967-977.

Wei, F., Liu, S., Song, Y., Pan, S., Zhou, M. X., Qian, W., Shi, L., Tan, L. and Zhang, Q. (2010) ‘TIARA: a visual exploratory text analytic system’, *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*. Washington, DC, USA, July 25-28, 2010. Association for Computing Machinery.

Wilensky, R. (1983) ‘Story grammars versus story points’, *Behavioral and Brain Sciences*, 6(4), pp. 579-591.

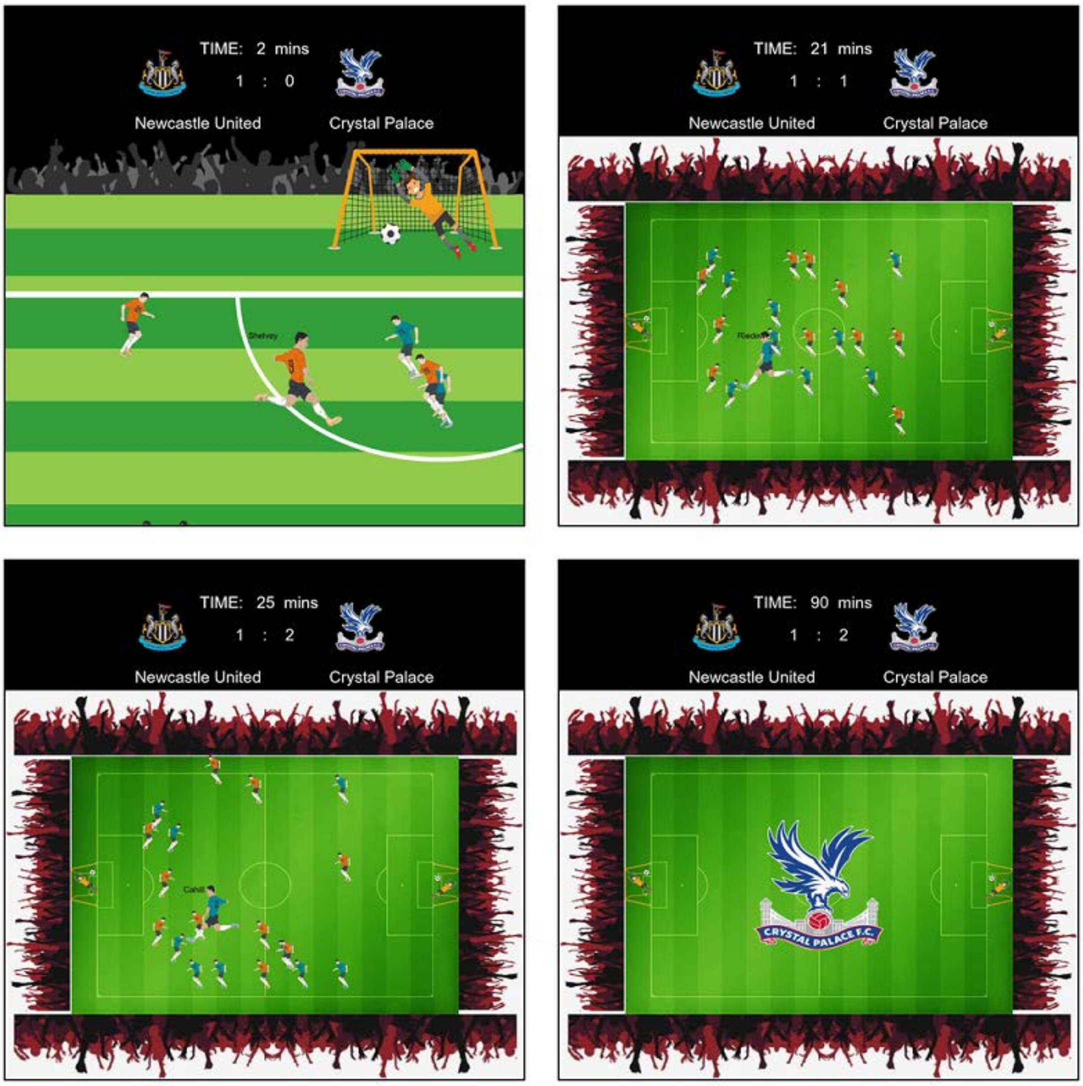
Wise, J. A., Thomas, J. J., Pennock, K., Lantrip, D., Pottier, M., Schur, A. and Crow, V. (1995) ‘Visualizing the non-visual: spatial analysis and interaction with information from text documents’, *Proceedings of Visualization 1995 Conference*. Atlanta, GA, USA, 30-31 Oct. 1995. IEEE.

Yang, G., Wen, D., Kinshuk, Chen, N. S. and Sutinen, E. (2012) ‘Personalized Text Content Summarizer for Mobile Learning: An Automatic Text Summarization System with Relevance Based Language Model’, *2012 IEEE Fourth International Conference on Technology for Education*. Hyderabad, India, 18-20 July 2012. IEEE.

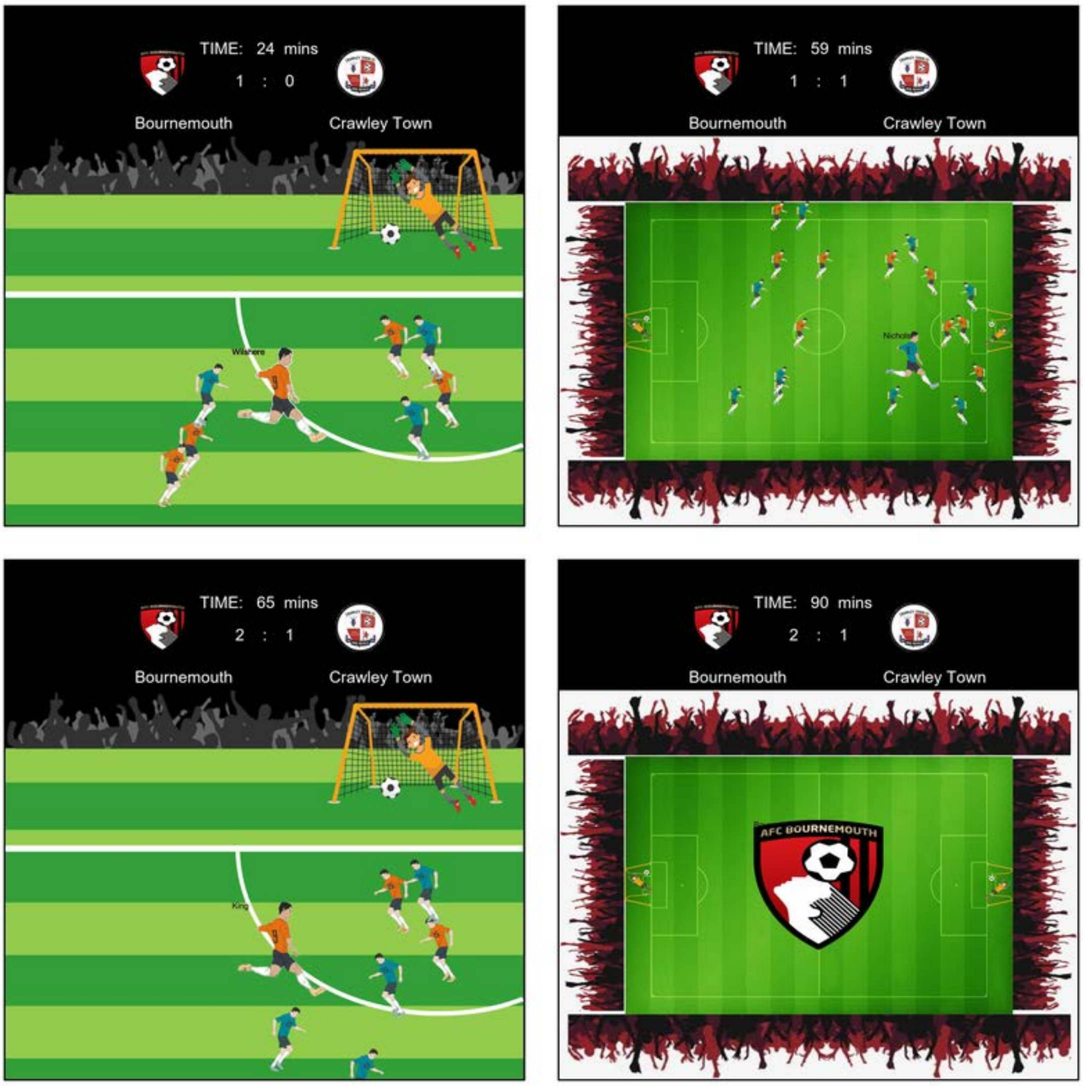
Zechner, K. (1996) ‘Fast generation of abstracts from general domain text corpora by extracting relevant sentences’, *Proceedings of the 16th conference on Computational linguistics-Volume 2*. Copenhagen, Denmark, August 5-9, 1996. Association for Computational Linguistics.

Zhang, P. Y. and Li, C. H. (2009) ‘Automatic text summarization based on sentences clustering and extraction’, 2009 2nd IEEE International Conference on Computer Science and Information Technology. Beijing, China, 8-11 Aug. 2009. IEEE.

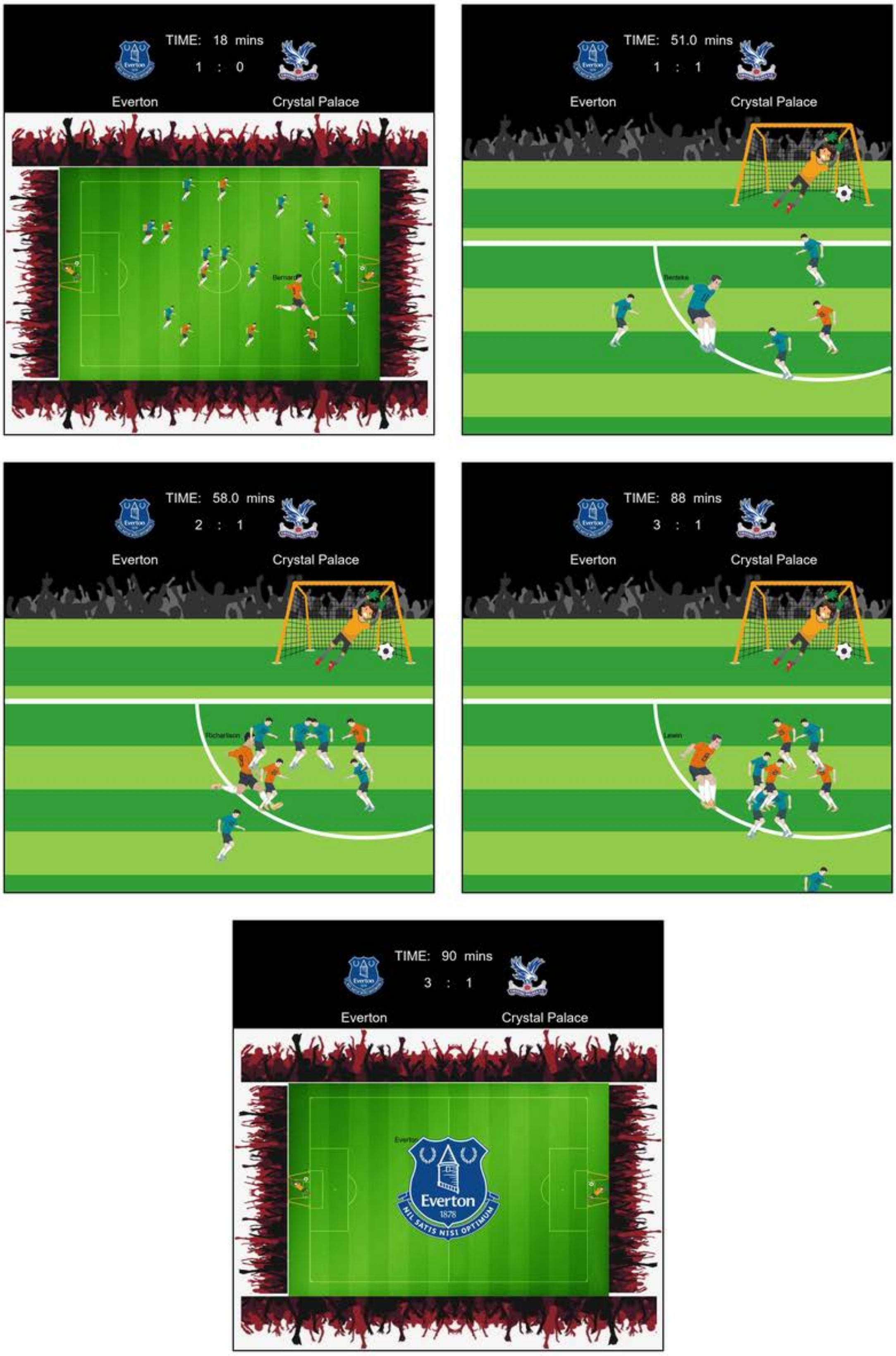
Appendix



Comic for Newcastle United vs Crystal Palace



Comic for Bournemouth vs Crawley Town



Comic for Everton vs Crystal Palace

Comic for Brora vs Hearts

Comic for Blackpool vs Peterborough United

Test Statistics

N	24
Kendall's W[a]	.652
Chi-Square	125.278
df	8
Asymp. Sig.	.000

a. Kendall's Coefficient of Concordance

Non-parametric test for Total Time in web experiment

Test Statistics

N	24
Kendall's W[a]	.712
Chi-Square	136.767
df	8
Asymp. Sig.	.000

a. Kendall's Coefficient of Concordance

Non-parametric test for Reading Time in web experiment

Test Statistics

N	24
Kendall's W[a]	.092
Chi-Square	17.620
df	8
Asymp. Sig.	.024

a. Kendall's Coefficient of Concordance

Non-parametric test for Answering Time in web experiment

Test Statistics

N	24
Kendall's W[a]	.458
Chi-Square	87.959
df	8
Asymp. Sig.	.000

a. Kendall's Coefficient of Concordance

Non-parametric test for Correct Rate in web experiment

Test Statistics[a]

N	11
Chi-Square	54.451
df	8
Asymp. Sig.	.000

a. Friedman Test

Non-parametric test for Agreement Rate in web experiment

Test Statistics[a]

N	24
Chi-Square	132.435
df	8
Asymp. Sig.	.000

a. Friedman Test

Non-parametric test for Question 1 in web experiment

Test Statistics[a]

N	24
Chi-Square	75.904
df	8
Asymp. Sig.	.000

a. Friedman Test

Non-parametric test for Question 2 in web experiment

Test Statistics[a]

N	24
Chi-Square	122.278
df	8
Asymp. Sig.	.000

a. Friedman Test

Non-parametric test for Question 3 in web experiment

Paired Wilcoxon test for total time for web experiment

	S1O	S2O	S3O	S1S	S2S	S3S	S1C	S2C	S3C
S1O		0.607	0.230	0.000	0.000	0.000	0.000	0.000	0.000
S2O	0.607		0.043	0.000	0.000	0.000	0.000	0.000	0.000
S3O	0.230	0.043		0.000	0.000	0.000	0.000	0.000	0.000
S1S	0.000	0.000	0.000		0.241	0.440	0.977	0.052	0.391
S2S	0.000	0.000	0.000	0.241		0.030	0.265	0.179	0.775
S3S	0.000	0.000	0.000	0.440	0.030		0.424	0.002	0.052
S1C	0.000	0.000	0.000	0.977	0.265	0.424		0.052	0.440
S2C	0.000	0.000	0.000	0.052	0.179	0.002	0.052		0.097
S3C	0.000	0.000	0.000	0.391	0.775	0.052	0.440	0.097	

COMIC

Participant	Total Time	Time per Word	Total Fixations	Time per Fixation	Total Questions Correc
1	57.25	0.923387	185	0.309459	71.43%
2	56.944	0.918452	92	0.618957	71.43%
3	22.782	0.367452	81	0.281259	71.43%
4	50.82	0.819677	105	0.484	100.00%
5	37.777	0.609306	138	0.273746	71.43%
6	28.343	0.457145	96	0.29524	85.71%
7	81.019	1.306758	54	1.500352	71.43%
8	43.571	0.702758	142	0.306838	100.00%
9	28.345	0.457177	66	0.42947	85.71%
10	25.159	0.40579	88	0.285898	100.00%
11	30.263	0.488113	115	0.263157	85.71%
12	19.755	0.318629	16	1.234688	85.71%

Paired Wilcoxon test for Reading time for web experiment

	S1O	S2O	S3O	S1S	S2S	S3S	S1C	S2C	S3C
S1O		0.317	0.304	0.000	0.000	0.000	0.000	0.000	0.000
S2O	0.317		0.022	0.000	0.000	0.000	0.000	0.000	0.000
S3O	0.304	0.022		0.000	0.000	0.000	0.000	0.000	0.000
S1S	0.000	0.000	0.000		0.043	0.230	0.317	0.007	0.011
S2S	0.000	0.000	0.000	0.043		0.019	0.012	0.407	0.475
S3S	0.000	0.000	0.000	0.230	0.019		0.864	0.002	0.008
S1C	0.000	0.000	0.000	0.317	0.012	0.864		0.007	0.007
S2C	0.000	0.000	0.000	0.007	0.407	0.002	0.007		0.886
S3C	0.000	0.000	0.000	0.011	0.475	0.008	0.007	0.886	

Paired Wilcoxon test for Answering time for web experiment

	S1O	S2O	S3O	S1S	S2S	S3S	S1C	S2C	S3C
S1O		0.179	0.265	0.819	0.841	0.219	0.587	0.627	0.241
S2O	0.179		0.710	0.110	0.189	0.171	0.007	0.040	0.100
S3O	0.265	0.710		0.162	0.346	0.909	0.056	0.056	0.977
S1S	0.819	0.110	0.162		0.954	0.331	0.440	0.587	0.049
S2S	0.841	0.189	0.346	0.954		0.179	0.732	0.587	0.199
S3S	0.219	0.171	0.909	0.331	0.179		0.067	0.116	0.607
S1C	0.587	0.007	0.056	0.440	0.732	0.067		0.841	0.043
S2C	0.627	0.040	0.056	0.587	0.587	0.116	0.841		0.086
S3C	0.241	0.100	0.977	0.049	0.199	0.607	0.043	0.086	

Paired Wilcoxon test for Agreement Rate for web experiment

	S1O	S2O	S3O	S1S	S2S	S3S	S1C	S2C	S3C
S1O		0.591	0.138	0.473	0.385	0.011	0.005	0.005	0.005
S2O	0.591		0.540	0.235	0.833	0.053	0.005	0.007	0.005
S3O	0.138	0.540		0.138	0.476	0.108	0.008	0.012	0.003
S1S	0.473	0.235	0.138		0.161	0.012	0.011	0.011	0.034
S2S	0.385	0.833	0.476	0.161		0.092	0.004	0.005	0.007
S3S	0.011	0.053	0.108	0.012	0.092		0.003	0.003	0.003
S1C	0.005	0.005	0.008	0.011	0.004	0.003		0.083	0.916
S2C	0.005	0.007	0.012	0.011	0.005	0.003	0.083		0.892
S3C	0.005	0.005	0.003	0.034	0.007	0.003	0.916	0.892	

Paired Wilcoxon test for Question 1 for web experiment

	S1O	S2O	S3O	S1S	S2S	S3S	S1C	S2C	S3C
S1O		1.000	0.285	0.000	0.000	0.016	0.000	0.000	0.000
S2O	1.000		0.248	0.000	0.000	0.005	0.000	0.000	0.000
S3O	0.285	0.248		0.000	0.000	0.000	0.000	0.000	0.001
S1S	0.000	0.000	0.000		0.439	0.066	0.008	0.008	0.008
S2S	0.000	0.000	0.000	0.439		0.005	0.008	0.008	0.008
S3S	0.016	0.005	0.000	0.066	0.005		0.000	0.000	0.000
S1C	0.000	0.000	0.000	0.008	0.008	0.000		1.000	1.000
S2C	0.000	0.000	0.000	0.008	0.008	0.000	1.000		1.000
S3C	0.000	0.000	0.001	0.008	0.008	0.000	1.000	1.000	

Paired Wilcoxon test for Question 2 for web experiment

	S1O	S2O	S3O	S1S	S2S	S3S	S1C	S2C	S3C
S1O		0.134	0.822	0.022	0.006	0.000	0.197	0.046	0.023
S2O	0.134		0.265	0.122	0.066	0.006	0.022	0.002	0.002
S3O	0.822	0.265		0.007	0.004	0.001	0.083	0.007	0.007
S1S	0.022	0.122	0.007		0.669	0.160	0.001	0.001	0.001
S2S	0.006	0.066	0.004	0.669		0.204	0.000	0.000	0.000
S3S	0.000	0.006	0.001	0.160	0.204		0.000	0.000	0.000
S1C	0.197	0.022	0.083	0.001	0.000	0.000		0.180	0.257
S2C	0.046	0.002	0.007	0.001	0.000	0.000	0.180		1.000
S3C	0.023	0.002	0.007	0.001	0.000	0.000	0.257	1.000	

Paired Wilcoxon test for Question 3 for web experiment

	S1O	S2O	S3O	S1S	S2S	S3S	S1C	S2C	S3C
S1O		0.366	0.405	0.381	0.134	1.000	0.000	0.000	0.000
S2O	0.366		0.000	0.090	0.000	0.000	0.000	0.000	0.000
S3O	0.405	0.000		0.167	0.057	0.000	0.000	1.000	0.000
S1S	0.381	0.090	0.167		0.599	0.415	0.000	0.000	0.000
S2S	0.134	0.000	0.057	0.599		0.180	0.000	0.053	0.000
S3S	1.000	0.000	0.000	0.415	0.180		0.000	0.439	0.465
S1C	0.000	0.000	0.000	0.000	0.000	0.000		1.000	0.317
S2C	0.000	0.000	1.000	0.000	0.053	0.439	1.000		0.317
S3C	0.000	0.000	0.000	0.000	0.000	0.465	0.317	0.317	