Educational Psychology Review (2025) 37:25
https://doi.org/10.1007/s10648-025-10000-y

REFLECTION ON THE FIELD

How Eye Read: A Social Network Approach

Leen Catrysse[1] · Tine van Daal[2] · Halszka Jarodzka[1] · Johanna K. Kaakinen[3] · Vincent Donche[2] · David Gijbels[2]

Accepted: 18 February 2025 / Published online: 13 March 2025

Abstract

The aim of the current paper is to offer a unique perspective on eye movement analysis in reading research by applying techniques from social network analysis to examine integration processes between sentences during reading. In a first step, we explored how network measures relate to the often-used duration measures in reading research in order to examine whether there is an additional value in using network measures. In a second step, we further explored how differences in network measures are related to text (i.e., topic structure) and reader characteristics (i.e., WMC). Thirty-one participants read three short expository texts. Four network measures at the sentence level were calculated for the three texts: strength, betweenness centrality, harmonic centrality, and local clustering coefficient. Correlations were computed between first-pass reading time and second-pass reading time and the network measures. Network measures were analyzed with (generalized) linear mixed-effects models. The results show that strength is strongly correlated to second-pass reading time. Betweenness, harmonic centrality, and the local clustering coefficient are not related to these often-used duration measures and thus capture aspects of integration processes that cannot be captured with duration measures. The results demonstrated that strength and betweenness centrality are related to reader's WMC. It was also shown that strength, harmonic centrality, and local clustering coefficient were related to the topic structure of the text. This study demonstrates that a social network approach offers a novel perspective on moment-to-moment integration processes during reading.

Keywords Reading processes · Social network analysis · Eye movements

✉ Leen Catrysse
leen.catrysse@ou.nl

[1] Department of Online Learning and Instruction, Faculty of Educational Sciences, Open Universiteit, Valkenburgerweg 177, Heerlen, the Netherlands

[2] Department of Training and Education Sciences, Faculty of Social Sciences, University of Antwerp, Antwerp, Belgium

[3] University of Turku, Turku, Finland

Introduction

Written texts are an important medium through which readers acquire new knowledge. Thus, reading is a crucial skill throughout school and adult life (Kendeou & O'Brien, 2018; Kendeou & Trevors, 2012; McNamara, 2011). To truly comprehend texts, readers must actively engage with the text to construct a mental representation of the text's content (Fox, 2009; Kintsch, 1988, 1998). A successful construction of a mental representation relies on processes associated with understanding sentences and the relation between sentences for which integration between sentences is crucial (McNamara & Magliano, 2009).

Individual differences in how mental representations are exactly built are tremendous (Moss et al., 2023). To gain more insight into how readers construct these mental representations, researchers have examined readers' eye movements during reading as they can reveal important individual differences in moment-to-moment processes during reading (Hyönä et al., 2003; Jarodzka & Brand-Gruwel, 2017; Rayner, 1998). These differences can be linked to characteristics of the reader and/or the text. Regarding reader characteristics, research demonstrated that working memory capacity, text-based interest, and prior knowledge among others influence eye movements (Catrysse et al., 2018; Jarodzka & Brand-Gruwel, 2017; Kaakinen et al., 2002; Moss et al., 2023). Other eye tracking studies showed that eye tracking measures may also vary according to characteristics of the sentences that are processed. Research has focused on processing differences between topic-introducing, topic-medial, and topic-final sentences (Ariasi et al., 2017; Hyönä & Lorch, 2004), relevant and irrelevant sentences (Kaakinen & Hyönä, 2005, 2007; Kaakinen et al., 2002), and central versus peripheral ideas in the text (Yeari et al., 2016).

Current eye tracking studies mainly focused on how much time is spent on sentences in a text (Hyönä et al., 2003; Jarodzka & Brand-Gruwel, 2017; Kaakinen, 2017). However, Hyönä and colleagues (2003) also indicated that transitions between sentences can provide crucial information on global text processing, which involves linking together information from sentences that are not adjacent in the text (Hyönä et al., 2003). Even though, the duration measures proposed by Hyönä and et al. (2003) include these important transitions from one sentence to another; they do not allow examining these transitions between different parts of the text. Transition matrices include all between-sentence movements and their frequencies and allow examining integration between non-adjacent parts in the text; however these matrices might become too complex for statistical analysis (Hyönä et al., 2003). Here, social network analysis can be invaluable, capturing information about those transitions into structural relations in network metrics and allowing analyzing important transitions between sentences. In addition, these network metrics may also provide insights into how information is eventually structured in the reader's mental representation (Ma et al., 2023).

Theoretical models emphasize the importance of sentence integration in reading processes (McNamara & Magliano, 2009) and representing mental representations of a text as networks of nodes (words, propositions, and sentences) and links (relationships between nodes) aligns with these models (Kintsch, 1998; McNamara &

Magliano, 2009; van den Broek, 2010). However, empirical studies using network science to study reading processes are scarce (Ma et al., 2023). Castro and Siew (2020) highlight the potential of network analysis for understanding complex systems like mental representations and their influence on reading processes. Social network analysis offers well-established quantitative measures to study these processes (Siew et al., 2019) and network metrics can capture mental processes that are not easily observable in other eye movement measures (Ma et al., 2023). Therefore, this study aims to use social network metrics to better understand the integration processes during reading and to provide a solution to analyze transitions between sentences for longer texts.

Theoretical Framework

Construction of Mental Representations

One of the most complete and influential theories of reading comprehension is the construction-integration (CI) model (Kendeou & O'Brien, 2018; Kintsch, 1998; McNamara & Magliano, 2009). In the CI model, reading comprehension requires building up a situation model of the text (Kintsch, 1998). This situation model is built by iterative processes in which current text is integrated with earlier textual input or with prior knowledge (Kintsch, 1998). First, in the construction phase, information can be activated from the current sentence, the prior sentence or text, and from prior knowledge (Kendeou & O'Brien, 2018; McNamara & Magliano, 2009). This activated textual information is then connected into a network of propositions (Kendeou & O'Brien, 2018). Textual information is iteratively integrated into the network and only those nodes are kept in the network that is connected to many others in the network. The integration of current textual input with prior text and prior knowledge is referred to as *inferencing* (McNamara & Magliano, 2009). Such inferences provide connections between the current read text, the previously read text, and prior knowledge (Clinton & van den Broek, 2012). Generating inferences is crucial when reading because (more complex) texts do not state every single relation between ideas presented within a text (Kintsch, 1998).

One important factor that has been examined in terms of individual differences in reading, with pervasive effects on the mental representation, is the global processing strategy adopted by the reader (Hyönä et al., 2002; Zhang et al., 2019). One way to investigate this global processing strategy is by examining the readers' processing of sentences relevant to the topic structure (Hyönä et al., 2002). In order to represent the topic structure of a text, the reader needs to be alert to the introduction of new subtopics and their relation with the overarching topic of the text (Hyönä et al., 2002). Some readers mention to look for topic sentences and skilled readers also focus on the topic-ending sentences (Pressley & Afflerbach, 1995). Successful readers thus devote extra attention at junctures in a text that represent important transitions in the topic structure such as topic-introducing and topic-ending sentences (Hyönä et al., 2002). Research has found that when paying extra attention at topic-introducing and -ending sentences, readers integrate their evolving mental representation of the text (Hyönä & Lorch, 2004; Hyönä et al., 2002).

Another important extensively studied individual reader characteristic in reading is working memory capacity (WMC) and is related to the kind of global processing strategies that reader use (Daneman & Carpenter, 1980; Daneman & Merikle, 1996; Moss et al., 2023; Unsworth & McMillan, 2013). WMC is needed to actively keep prior text information activated while reading new sentences, so the reader is able to integrate prior and current information in the text, to activate relevant information from long-term memory, and to maintain attentional focus on the reading task (Daneman & Carpenter, 1980). In addition, WMC is also needed to make strategic decisions during reading such as rereading to build a high-quality mental representation (Daneman & Merikle, 1996). Readers with lower WMC have less capacity to integrate information from the text and prior knowledge into their mental representation (Daneman & Carpenter, 1980; Daneman & Merikle, 1996). Readers with a high working memory capacity are more strategic in allocating their attention and might use a more selective reading strategy (Kaakinen et al., 2002; Moss et al., 2023), are better in drawing inferences (Linderholm, 2002; Singer et al., 1992), and increase attention to important text areas (Moss et al., 2023).

Using Eye Tracking to Uncover Differences in Integration Processes

Eye tracking allows us to investigate at which parts of the text someone looks, for how long and in which order (Holmqvist et al., 2011; van Gog & Jarodzka, 2013), providing detailed information about the reading process. In reading, two types of eye movements are most important, namely fixations and saccades (Rayner, 1998, 2009). During fixations, the eye is almost still and information can be extracted from the text. During saccades, on the other hand, the visual focus of attention is relocated to another place in the text. While the eye is moving rapidly when executing a saccade, we are not able to extract any information from a text (Holmqvist et al., 2011; Rayner, 1998, 2009). Fixations, as indications of information intake, are usually related to and summarized across specific areas of interest (AOIs). In reading research, this can be a character, a word, a phrase, or a sentence. Fixations within one AOI, like a sentence, are subsequently aggregated to calculate the first-pass fixation duration, second-pass fixation duration or total fixation durations.

First-pass fixation duration refers to the duration of all fixations when initially encountering a target region (e.g., a sentence). Second-pass fixation durations refer to the duration of all regressive fixations back to the target region after the first-pass reading has been terminated (Hyönä et al., 2003). First-pass fixation durations are an indication of early processing. Second-pass reading times reflect processes happening later in comprehension, such as cognitive conflict (Mikkilä-Erdmann et al., 2008), high-level or deeper cognitive processing (Ariasi & Mason, 2011; Catrysse et al., 2018; Penttinen et al., 2012), comprehension monitoring (van Gog & Jarodzka, 2013), difficult passages (Rayner, 2009), and attempts to reinstate text information into working memory in order to elaborate on it or rehearse it (Hyönä & Lorch, 2004). Readers are not usually aware of the time they have spent during first-pass reading of sentences, but they can accurately report which parts of text they have look-backed to perform second-pass reading (Hyönä & Nurminen, 2006). This indicates that second-pass reading times reflect strategic processing of text.

In eye tracking research, shifts between AOIs are called transitions (Holmqvist et al., 2011). These transitions can be linked to important integration processes during reading (Alemdag & Cagiltay, 2018). Transitions are often represented in matrices, where each row and each column of the matrix stands for one sentence of the text and the cells contain the number of transitions between sentences (Holmqvist et al., 2011). Research has paid little attention to how readers transition between different parts of a text. One notable exception is the use of scan path analyses, which examine how readers visually explore various sections of the text and in what sequence (Mézière et al., 2024; von der Malsburg & Vasishth, 2011). These analyses are effective in identifying common patterns of eye movements that are linked to, for example, successful reading comprehension (Mézière et al., 2024). Another promising method for studying these transitions is social network analysis (Ma et al., 2023), which captures the connections and integration between different text sections, something that cannot be achieved by scan path analyses.

Social Network Analysis to Unravel Integration Processes

A social network consists of a set of nodes connected with some type of relations to each other, which are often called edges in graph theory (Borgatti, 2005; Borgatti et al., 2009, 2013; Siew et al., 2019). In these networks, definitions of nodes and edges depend on the cognitive phenomena that are examined (Ma et al., 2023). A few studies have applied network metrics to analyze eye tracking data in which different AOIs are used as nodes and the transitions between AOIs as edges (Ma et al., 2023; Zhu & Feng, 2015). Zhu and Feng (2015) used social network analysis to examine individual differences in mathematical problem solving. Eye tracking data was collected while students were solving a math problem. Eleven AOIs were identified such as the questions, the graph, the image, and corresponding question response areas. A transition matrix was created that summarizes the transitions between AOIs. Their work demonstrated that high-performing students made more strategic transitions between AOI triplets compared to low-performing students, as indicated by several triad structures. A recent study of Ma et al. (2023) has employed a social network approach to analyze eye tracking data while reading an expository text. In their study, participants read sentences one by one. Nodes in the networks were words, and edges in the network were transitions between words in a single sentence. Their results demonstrated that low and high ability readers network metrics showed distinctive differences, which are reflected in network metrics such as density, centrality, small-worldness, transitivity, and global efficiency. Skilled readers' reading networks were characterized by a lower density, transitivity and global efficiency, and a higher centrality compared to less skilled readers. Ma et al. (2023) indicated that these patterns are consistent with previous eye tracking studies that has shown that skilled readers make fewer fixations, more skips, and fewer regressions during reading.

There are several types of structures in a social network: the microscopic, mesoscopic, and macroscopic structure (Siew, 2020; Siew et al., 2019). The microscopic structure refers to structural properties of individual nodes and edges. The mesoscopic structure refers to a subset of nodes and substructures. The macroscopic structure summarizes the entire network structure. Previous studies using network analysis for eye tracking data focused on the macroscopic network structure (Ma et al., 2023; Zhu & Feng, 2015). Since the integration between sentences is crucial for building a situation model (McNamara & Magliano, 2009) and topic structure plays a role in this (Hyönä et al., 2002), this study will focus on *microscopic network measures*.

At the microscopic level, network analysis mostly focuses on quantifying the importance of a node in the graph via centrality measures (Borgatti, 2005; Borgatti et al., 2013; Siew, 2020; Siew et al., 2019). Three different but related centrality measures are defined in the network literature namely *strength*, *closeness*, and *betweenness* (Barrat et al., 2004; Freeman, 1978; Opsahl et al., 2010). When calculating these measures for reading networks, in which each sentence is a node and the number of transitions between sentences are edges, these metrics can indicate which sentences are more important in the network and how this is related to individual differences. All centrality measures can indicate whether sentences occupy central and important locations in the mental representations (Siew, 2020). The centrality measures are strongly related and the choice between centrality measures depends on specific characteristics of the network (Borgatti, 2005; Borgatti et al., 2013; Siew, 2020; Siew et al., 2019). Because research is lacking on centrality measures in reading research, it is not yet clear which measures are most suitable to investigate integration processes during reading.

Strength refers to the number of edges connected to a node (Siew et al., 2019) or in terms of reading it reflects the number of transitions connected to a particular sentence and thus how much a sentence has been visited. Nodes with a higher strength have an important role in exchanging information across the network (Siew et al., 2019). High-strength nodes are more likely to be accessed during reading, while low-strength nodes are accessed less frequently. High-strength nodes facilitate the integration of new information. In loosely connected networks, strength tends to be lower due to the sparse and weaker connections, while in highly connected networks, strength tends to be higher due to the denser and stronger connections between nodes. Imagine reading a paragraph with one topic-initial sentence (TI), three topic-medial sentences (M1-3), and one topic-final sentence (TF). In Fig. 1, two different hypothetical reading networks are visualized. The left part represents a very linear reading pattern in which the reader goes linearly through the sentences without going back. The right part represents a more integrated reading pattern in which the reader also makes connections from the topic-initial sentence to all other sentences. In the left part, node TI has a strength of 1 because there is only one transition between nodes TI and M1. In the right part, node TI has a strength of 4 because it is connected to nodes M1, M2, M3, and TF.

Closeness centrality indicates to what extent a node is close to other nodes in the network (Siew, 2020) and is calculated as the inverse of the shortest path between nodes. It quantifies the proximity of a sentence to all other sentences. A sentence

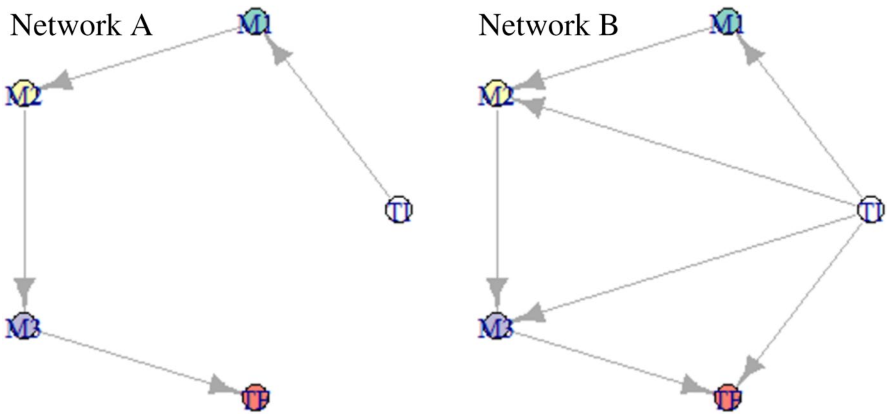

Fig. 1 Example networks

with high closeness centrality is close to many other sentences in the network since there is a short path between that sentence and other sentences. A sentence with low closeness centrality is far from many other sentences, as there is a long path between that sentence and the other sentences. In terms of reading, a sentence with high closeness centrality is characterized by transitions to many other sentences in the reading network and indicates more integration between that sentence and other sentences. High closeness centrality nodes serve as key integrators within the reading network because they help in connecting various sentences in the reading network. Closeness centrality reflects how readily information can spread through a network, with nodes in highly connected networks typically having higher closeness centrality due to the shorter average path lengths. In the left part of Fig. 1, the shortest path from TI to M1 is 1, from TI to M2 is 2, from TI to M3 is 3, and from TI to TF is 4. The sum of these shortest paths is 10, and the inverse of 10 is the average closeness centrality for node TI, which is 0.10. In the right part of Fig. 1, the shortest path from TI to M1 is 1, from TI to M2 is 1, from TI to M3 is 1, and from TI to TF is 1. The average closeness centrality for TI is 0.25 (the inverse of 4). Sentences with the highest closeness centrality are considered as more central in the reading network. Since node TI is connected to more sentences in the right part of Fig. 1, it is more central and receives a higher closeness centrality score.

Betweenness centrality measures the extent to which a node lies on the shortest path between any two nodes in the network (Siew, 2020). A node with high betweenness centrality frequently lies in between the shortest path between all possible pairs of nodes, while a node with low betweenness centrality does not tend to lie on the shortest path between pairs of nodes. Nodes with high betweenness centralities might represent landmark concepts that are crucial linking concepts (Siew, 2020) and bridge distant sentences in the reading network. In terms of reading, sentences with a high betweenness centrality are often used as an anchor to integrate information between two other sentences. High betweenness centrality nodes serve as crucial intermediaries connecting different regions within the reading network. When calculating betweenness centrality for node TI in the left part of Fig. 1, we need to consider all pairs of nodes, not including node TI: M1 → M2, M1 → M3, M1 → TF, M2 → M3, M2 → TF, and M3 → TF. Looking at the shortest paths between these pairs, node TI is never on the shortest path between the pairs and thus betweenness centrality for node TI is 0. In the right part of Fig. 1, the same pairs can be considered for node TI. Here again, node TI is never on the shortest path between the pairs (taking directions of the arrows into account) and also receives a betweenness centrality score of 0. This demonstrates that betweenness centrality emphasizes a different aspect of integration compared to strength and closeness centrality.

Another network measure that is related to reading is the *local clustering coefficient*, which specifies to what extent the neighbors of a node are interconnected (Siew et al., 2019). Where centrality metrics assess the node's role and importance in the network, the local clustering coefficient focuses on the local neighborhood of a node and measures the connectivity among its immediate neighbors. The local clustering coefficient represents the likelihood for two nodes to be connected when they share a mutual neighbor (Siew et al., 2019). In terms of reading, a high local clustering coefficient is related to a high number of transitions between neighboring sentences of a particular sentence (Ma et al., 2023). High local clustering coefficient nodes indicate intense local processing of information around that node. It is important to note that neighboring sentences refer to sentences that a given sentence is connected to and might not be equal to the preceding and following sentence. In the left part of Fig. 1, node M1 has two neighboring nodes to which it has connections: TI and M2. Of all possible connections between the neighbors of M1 there is no connection, and thus M1 receives a local clustering coefficient score of 0. In the right part of Fig. 1, node M1 has the same neighboring nodes, but they are connected to each other (TI-M2). Therefore node M1 receives a local clustering coefficient score of 1 (connected pairs/total amount of possible connections). In this scenario, there are also transitions between the neighboring sentences of M1.

The different network measures do resemble each other but emphasize different kinds of importance in a network. The different centrality measures for the networks visualized in Fig. 1 are presented in Table 1. For network A (left part of

Table 1 Centrality measures for network examples A and B

	TI	M1	M2	M3	TF
Network A					
Strength	1	2	2	2	1
Betweenness	0	3	4	3	0
Closeness	0.10	0.14	0.17	0.14	0.10
Local clustering coefficient	0	0	0	0	0
Network B					
Strength	4	2	3	3	2
Betweenness	0	0	0.17	0.17	0
Closeness	0.25	0.17	0.20	0.20	0.17
Local clustering coefficient	0.50	1	0.67	0.67	1

Fig. 1), the second medial sentence (M2) receives the highest score for strength, closeness centrality, and betweenness centrality. For network B (right part of Fig. 1), the topic-initial sentences (TI) has the highest strength since this sentence has the most connections or transitions. However, the second (M2) and third medial sentence (M3) show the highest betweenness centrality and thus often lie on the shortest path between two other nodes in the network. These sentences represent crucial linking sentences. This example already demonstrates how strength and betweenness centrality offer a different view on the importance of a node in the network. The topic-introducing sentence has the highest closeness centrality. The second (M2) and third medial sentence (M3) show the highest local clustering coefficient, indicating that there is more integration between the neighbors of these sentences.

This Study

The goal of this study is to apply social network analysis on eye movement data from reading to gain insight into which microscopic network measures can be related to integration processes during reading. In a first step, we explored how centrality measures from social network science relate to the often-used duration measures in reading research. Ma et al. (2023) indicated that network metrics can capture mental processes that are not easily observable in eye movement duration measures. By examining the relations between duration measures and network metrics, we explicitly examine what the added value is of these network measures and whether they indeed do capture mental processes that cannot be captured with duration measures.

RQ1: How are microscopic network measures related to first-pass and second-pass duration measures?

In a second step, we further explored how differences in network measures are related to both text and reader characteristics. Since both the topic structure (Hyönä & Lorch, 2004; Hyönä et al., 2002) and WMC (Daneman & Carpenter, 1980; Daneman & Merikle, 1996; Moss et al., 2023; Unsworth & McMillan, 2013) affect integration processes during reading, it is examined whether they are related to differences in network measures. Centrality measures can reflect structural characteristics of the text (Ma et al., 2023). Ma et al. (2023) describe that reading networks of expository texts will usually be more centralized than reading networks of narrative texts. This is because expository texts often illustrate specific topics by providing relative descriptions, while narratives tend to follow a timeline which leads to a more linear structure. It can be assumed that this centrality measures can capture other structural differences, such as the topic structure of texts, but so far this has not yet been examined.

RQ2: How are microscopic network measures related to topic structure and WMC?

Methodology

The dataset for this study was reused from an already published study (Catrysse et al., 2018). However, a new analytic approach was used to answer different research questions. The dataset and scripts are available at osf.io (https://osf.io/qbctd/).

Participants

The participants for the study were higher education freshmen in social sciences. A sample of 42 students took part in the study. Due to calibration problems and common problems with eye tracking data quality (Holmqvist et al., 2011), data of 31 participants with a mean age of 20.23 ($SD = 2.12$) was available for the analyses. The calibration accuracy was visually inspected together with the replay of eye movements on the text. When the calibration was not successful, or drift occurred, eye movement data were removed from the dataset. The participants received two cinema tickets, and informed consent was obtained. All participants had normal or corrected-to-normal vision, reported having no learning disorders, and Dutch was their native language.

Materials and Instruments

Texts

Students were asked to study three expository texts on positive psychology in Dutch. Positive psychology was not included in their curriculum, so participants had little prior knowledge on the topic. Texts were adapted from the World Book of Hope (Bormans, 2016). The topics of the texts were on hope and happiness (414 words, 2161 characters, and 22 sentences), the tyranny of positive thinking (386 words, 1986 characters, and 18 sentences) and music and hope (392 words, 2076 characters, and 19 sentences). Each text consisted of four paragraphs. For the text on hope and happiness, the first paragraph is about the definition and importance of hope, the second about measuring hope, the third about the definition and measurement of happiness, and the fourth about the relationship between hope and happiness. For the text on the tyranny of positive thinking, the first paragraph is about the philosophy of positive thinking, the second about research on hope and healing, the third on challenges with positive psychology programs, and the fourth on limits of positive thinking. For the text on music and hope, the first paragraph dealt with implicit knowledge of music, the second with music and emotions, the third with personal preferences, and responses to music and the fourth with research on music and hope. All texts thus consisted of subtopics that are related to each other and can be found on the OSF project under "(1) Materials." The three texts were similar on lexical and sentence complexity, as checked with T-Scan (Pander Maat et al., 2014).

Eye Tracking Equipment

The Tobii TX300 eye tracker (dark pupil tracking) was used to collect students' eye movements. The eye-tracking component was integrated into a 23-inch TFT monitor with a maximum resolution of 1920 × 1080 pixels. The eye tracking camera sampled data binocularly at the rate of 300 Hz. A head stabilization system was not required, and head movement was allowed (37 × 17 cm). Tobii Technology (Stockholm, Sweden) reported a gaze accuracy of 0.4°, gaze precision of 0.15°, and latency between 1.0 and 3.3 ms for this eye tracker. The eye movements were recorded with Tobii-Studio (3.2) software.

Working Memory Capacity

Working memory capacity was measured by the automated operation span task (AOST; Unsworth et al., 2005). Participants needed to solve a series of mathematical operations while trying to memorize a set of unrelated letters. To make sure that participants were not only focused on remembering the letters, a 85% accuracy criterion was used for solving the mathematical problems (Unsworth et al., 2005). It has been shown that both verbal complex span tasks (such as the reading and listening span) and operation span tasks with numerical stimuli are related to reading comprehension (Daneman & Merikle, 1996; McVay & Kane, 2012). The AOST provides an absolute and a partial credit score. Since the partial credit score is preferred over the absolute score, the latter was used (Conway et al., 2005). The score for this working memory capacity test was standardized for further analysis.

Procedure

The study was conducted individually for each participant during a 1-h session. Students first did the AOST on a computer. Next, participants received written instructions on the screen for the reading task, where they were asked to study the texts as if they were preparing for their exams. After that, the eye tracker was calibrated using a nine-point calibration procedure in which students needed to track nine red calibration dots on a plain, gray background. Students were seated about 60 cm from the screen for the calibration, and the eye tracker was recalibrated before each text. Texts were presented one at a time on the screen, so scrolling was not needed. Studying was self-paced and the presentation order of the texts was counterbalanced across participants.

Analysis

Eye Tracking Data

We used the Tobii I-VT fixation filter for fixation identification, which is an implementation of a classification algorithm proposed by Olsson (2007). It uses a velocity threshold (35 pixels per window) and a distance threshold (35 pixels) (Olsen, 2012).

For each sentence in the text, an area of interest (AOI) was defined. A distinction was made between topic-introducing, topic-medial and topic-ending sentences (Hyönä et al., 2002). The topic-introducing sentence of the paragraph is the first sentence of a paragraph, the topic-ending sentences is the concluding sentence of a paragraph, and topic-medial sentences are all sentences in between. Text 1 contained 22 sentences (topic-introducing = 4, topic-medial = 14, and topic-ending = 4); text 2 consisted of 18 sentences (topic-introducing = 4, topic-medial = 10, and topic-ending = 4) and text 3 counted 19 sentences (topic-introducing = 4, topic-medial = 11, and topic-ending = 4).

Fixation data was exported from the Tobii Pro Studio software and contained a list of consecutive fixations for each participant for each text linked to the AOIs. An example of a few fixations for one participant is presented in Table A in Online Resource 2. A string variable was generated that included the AOI at which a student was looking at per fixation (Table B, Online Resource 2). In a next step, all fixations that fell outside AOIs were removed. Successive identical AOIs were deleted and were given a weight (Table C, Online Resource 2). When a participant fixated three times on the same AOI (i.e., sentence) in a row, these three rows in the dataset were removed by one row indicating that that AOI was fixated three times. AOI strings were converted to AOI transition matrices for further analysis (Table D, Online Resource 2). These matrices contain information from the source of the transition and the target to which the transition is going and how many times this transition occurred during reading.

First-pass and second-pass reading times were calculated for each AOI. First-pass reading time refers to the summed duration of all fixations on the sentence before moving on another sentence. Two measures of first-pass reading were calculated using the eyeRead R-package (Verhavert et al., 2020; version 0.0.4). First-pass forward reading is the sum of all durations that land on unread parts of a sentence during first-pass reading. First-pass regressive reading is the sum of all durations that land on already read parts of the sentence during first-pass rereading. Second-pass reading time refers to the duration of all regressions back to the sentence after first-pass reading time has terminated (Hyönä et al., 2003).

The datasets and script for the analysis in this section can be found in the OSF project under "(2) Eye Tracking Data."

Social Network Metrics

For each text and for each student a social network was constructed, resulting in 31 networks per text. Weighted social network objects were created based on an edge list and a node list with the package igraph (version 2.0.3; Csárdi et al., 2024) in RStudio (version 2023.12.1). The edge list contained all transitions indicating the source and target and the number (or weight) of transitions (see Table D, Online Resource 2). The node list contained the overview of AOIs per text per participant. Weighted social network objects are created since they also include information about the number of transitions. Strength, betweenness centrality, harmonic centrality, as a special form of closeness centrality, and local clustering coefficient were calculated for each node in the weighted networks. Harmonic centrality was chosen, since it allows calculating closeness centrality for disconnected graphs and is calculated as the sum of the inverse of the shortest paths between nodes. Strength, betweenness centrality, and harmonic centrality were computed for each sentence in each text with the igraph package (version 2.0.3; Csárdi et al., 2024). The local clustering coefficient was computed with the DirectedClustering package (version 0.1.1; Clemente & Grassi, 2018).

An example of two reading networks is provided in the supplementary information (Online Resource 1) in which gaze plots on the text; network visualizations and tables with network metrics and duration measures are provided. The datasets and script for the analysis in this section can be found in the OSF project under "(3) Social Network Metrics."

Statistical Analysis

Regarding the first research question, descriptives were computed for the network and duration measures. Additional histograms per variable that show the distribution of each dependent variable can be found in the OSF project under "(4) Statistical Analysis". Next, correlations were computed between first-pass forward reading time, first-pass regressive reading time, and second-pass reading time, and the network measures (strength, betweenness, harmonic centrality, and local clustering coefficient). Concerning the second research question, the network measures were analyzed using linear mixed-effects models. Mixed-effects models are statistical models that incorporate random and fixed effects and allow analyzing unbalanced datasets (Baayen, 2008; Baayen et al., 2008). Participants, sentences, and texts were considered as crossed random effects (Baayen, 2008; Baayen et al., 2008). In each model, sentence type (topic-introducing, topic-medial, and topic-ending) and working memory capacity were added as fixed effects. Additional post hoc comparisons were performed because we compared three sentence types with each other, the standard output of the mixed-effects model only compares the baseline category with the two other categories and does not show all possible comparisons. To obtain all pairwise comparisons the emmeans package was used (version 1.10.1; Lenth, 2022). The marginal and conditional explained variance of the models was reported using the performance package (version 0.11.0, Lüdecke et al., 2021).

The dependent variables were strength, betweenness centrality, harmonic centrality, and local clustering coefficient. Regarding strength, a generalized linear mixed-effects model was estimated using the Poisson distribution using the lme4 package (version 1.1–35.2, Bates et al., 2015). The Poisson distribution was chosen because strength refers to the number of transitions and is thus count data. The lmerTest package was used to obtain *p*-values (version 3.1–3; Kuznetsova et al., 2017). Concerning betweenness and harmonic centrality, linear mixed-effects models were used since the data was close to the normal distribution. For the local clustering coefficient, which is a value between 0 and 1, a generalized linear-mixed effects model was estimated using a zero-inflated beta distribution with the glmmTMB package (version 1.1.9; Brooks et al., 2017).

The datasets, script for the analysis and histograms in this section can be found in the OSF project under "(4) Statistical Analysis."

Results

RQ1: How are Microscopic Network Measures Related to First-Pass and Second-Pass Duration Measures?

The descriptive statistics of the network measures and duration measures are presented in Table 2. The effect sizes of Cohen (1988) are used to interpret the correlations. When the correlation coefficient is lower than 0.10, there is an absent correlation; between 0.10 and 0.29, a weak correlation; between 0.30 and 0.49, a moderate correlation and between 0.50 and 1, a strong correlation. The same cut-off points are used for negative values. There is a negative moderate correlation between strength and first-pass forward reading time and a negative weak correlation between strength and first-pass rereading time (Table 3).

There is a strong positive correlation between strength and second-pass reading time. Betweenness is not correlated with first-pass forward reading time, first-pass rereading time, or second-pass reading time. Harmonic centrality is not correlated with first-pass forward reading time or first-pass rereading time and only weakly correlated with second-pass reading time. Local clustering coefficient is weakly correlated to first-pass forward reading time and first-pass rereading time and not correlated with second-pass reading time.

Table 2 Descriptives

	Mean	*SD*	Min	Max
First-pass forward reading time (ms/char)	18.285	14.480	0	90.323
First-pass regressive reading time (ms/char)	8.441	11.746	0	92.903
Second-pass reading time (ms/char)	36.464	38.895	0	316.83
Strength	8.329	6.413	0	58
Betweenness centrality	76.656	56.344	0	259.667
Harmonic centrality	6.223	1.974	0	14.167
Local clustering coefficient	0.046	0.081	0	0.565

Table 3 Correlations between centrality measures and duration measures

	First-pass forward reading time	First-pass regressive reading time	Second-pass reading time
Strength	−0.258	−0.094	0.645
Betweenness	0.035	0.086	0.022
Harmonic centrality	−0.0006	0.065	0.111
Local clustering coefficient	−0.130	−0.095	0.064

How are Microscopic Network Measures Related to Text and Reader Characteristics?

Descriptive statistics are presented in Table 4 for the microscopic network measures. The output of the mixed-effects models is presented in Table 5 and the output of the multiple comparisons in Table 6.

Strength

The results demonstrate that strength is related to both text and reader characteristics. Post hoc tests (Table 6) indicate that topic-medial sentences have a higher strength compared to topic-initial ($p < 0.001$) and topic-final sentences ($p < 0.001$). This indicates that more transitions are connected to medial sentences. Regarding working memory capacity, the results show that with increasing working memory capacity, strength increases for all sentences (Table 5). Readers with higher WMC make more transitions between sentences. The variance (65.1%) in strength can be explained by the full model (both random and fixed effects), and 12.5% can be explained by the fixed effects only (sentence type and WMC).

Betweenness Centrality

The results show that betweenness centrality is not related to text characteristics, as no differences were found between sentence types (Table 5). Betweenness is related to working memory capacity: when working memory capacity increases, betweenness decreases. This indicates that with a higher working memory capacity, sentences do not lie that often on shortest paths between other sentences and that reading networks are not characterized by having crucial intermediaries connecting different regions within the reading network. The variance (26%) in betweenness centrality can be explained by the full model (both random and fixed effects), and 2.4% can be explained by the fixed effects only (sentence type and WMC).

Harmonic Centrality

Regarding harmonic centrality, post hoc results indicate that topic-medial sentences have a lower harmonic centrality compared to topic-final sentences ($p = 0.004$). It means that topic-final sentences are less close to other sentences

Table 4 Descriptive statistics (*mean* and *SD*) for microscopic network measures

	Strength	Betweenness centrality	Harmonic centrality	Local clustering coefficient
Topic-initial	6.973 (5.259)	67.887 (55.519)	6.319 (2.085)	0.053 (0.096)
Topic-medial	9.375 (6.952)	77.788 (56.735)	6.092 (1.901)	0.050 (0.081)
Topic-final	6.634 (5.082)	82.122 (55.171)	6.508 (2.037)	0.028 (0.060)

Table 5 Output of the best fitting mixed effects models

Random effects	Strength				Betweenness centrality				Harmonic centrality				Local clustering coefficient			
	Variance	*SD*			Variance	*SD*			Variance	*SD*			Variance	*SD*		
Student	0.169	0.411			228.7	15.12			0.960	0.980			0.006	0.240		
Text	0.002	0.040			132.8	11.53			0.414	0.644			0.000	0.001		
AOI	0.021	0.145			406.2	20.15			0.107	0.327			0.008	0.281		
Residual					2409.6	49.09			2.518	1.587						
Fixed effects	β	*SE*	z	$pr(>\|z\|)$	β	*SE*	t	$pr(>\|t\|)$	β	*SE*	t	$pr(>\|t\|)$	β	*SE*	z	$pr(>\|z\|)$
Intercept	1.785	0.090	19.757	< 0.001	82.122	9.591	8.563	< 0.001	6.508	0.430	15.141	< 0.001	− 2.524	0.114	− 22.203	< 0.001
Intro	0.053	0.066	0.802	0.422	− 14.235	8.981	− 1.585	0.119	− 0.189	0.177	− 1.069	0.290	0.330	0.140	2.362	0.018
Medial	0.341	0.054	6.365	< 0.001	− 5.588	7.369	− 0.758	0.452	− 0.491	0.145	− 3.378	0.001	0.343	0.115	2.974	0.003
WMC	0.014	0.007	1.971	0.049	− 0.709	0.279	− 2.543	0.017	0.024	0.017	1.413	0.168	0.006	0.005	1.357	0.175
Explained variance	Marginal R^2		Conditional R^2		Marginal R^2		Conditional R^2		Marginal R^2		Conditional R^2		Marginal R^2		Conditional R^2	
	0.125		0.651		0.024		0.260		0.026		0.387		0.009		0.064	

Table 6 Parameter estimates of the multiple comparisons of means for sentence type (post hoc Tukey contrast with Bonferroni correction)

	Strength				Betweenness				Harmonic centrality				Local clustering coefficient			
	B	*SE*	*z*	*p*	*B*	*SE*	*t*	*p*	*B*	*SE*	*t*	*p*	*B*	*SE*	*z*	*p*
Final — Intro	−0.053	0.065	−0.802	0.702	14.24	8.98	1.585	0.261	0.189	0.177	1.069	0.537	−0.330	0.140	−2.362	0.048
Final — Medial	−0.342	0.054	−6.365	<0.001	5.59	7.37	0.758	0.730	0.491	0.145	3.378	0.004	−0.343	0.115	−2.974	0.008
Intro — Medial	−0.289	0.053	−5.404	<0.001	−8.65	7.37	−1.173	0.474	0.301	0.145	2.076	0.105	−0.014	0.111	−0.123	0.992

compared to medial sentences and thus readers transition less often from topic-ending sentences to other parts of the text compared to medial sentences. There is no relation between working memory capacity and harmonic centrality. The variance (38.7%) in harmonic centrality can be explained by the full model (both random and fixed effects), and 2.6% can be explained by the fixed effects only (sentence type and WMC).

Local Clustering Coefficient

Post hoc tests indicate that topic-final sentences have a lower local clustering coefficient compared to topic-initial ($p = 0.048$) and topic-medial sentences ($p = 0.008$). This result indicates that less rereading is going on around the topic-final sentence compared to topic-initial and topic-medial sentences. The variance (6.4%) in local clustering coefficient can be explained by the full model (both random and fixed effects) and 0.09% can be explained by the fixed effects only (sentence type and WMC).

Discussion and Conclusion

The goal of this study was to see whether social network approach sheds new light on individual differences in integration processes when reading. Even though building a mental representation of a text has theoretically been compared to a network with nodes and edges (Kintsch, 1998; McNamara & Magliano, 2009; Broek, 2010), network science has rarely been used to examine individual differences in reading processes (Ma et al., 2023). The study of Ma et al. (2023) used a social network approach on eye movement data at the word level during reading in a sentence-by-sentence paradigm. Our study focuses on using a network approach on eye movement data at the sentence level during reading short expository texts, which allows examining global text processing that involves linking together information from sentences that are not adjacent in the text (Hyönä et al., 2003).

The first research question addressed how the network measures are related to the often-used duration measures. Only strength, referring to the number of transitions connected to a sentence, was strongly correlated to second-pass reading time. Since second-pass reading time refers to the duration of all regressions back to a sentence (Hyönä et al., 2003), it makes sense that it is highly related to strength which is an indication of the number of transitions related to a sentence. Betweenness centrality, harmonic centrality and the local clustering coefficient are not related to these often-used duration measures and thus capture different processes. As already indicated by Ma et al. (2023), these network measures are able to capture mental processes associated with integration information in the text that cannot be captured by duration measures alone. Second-pass reading times are an indication of strategic processing and can reflect high-level or deeper cognitive processing (Ariasi & Mason, 2011; Catrysse et al., 2018; Penttinen et al., 2012) and attempts to reinstate text information into working memory in order to elaborate on it or rehearse it (Hyönä & Lorch, 2004). Network metrics, on the other hand, can reveal whether sentences occupy central and important locations in the reading networks (Siew, 2020). They are more related to structural properties of the reading network based on transitions and connections that are made between sentences. The microscopic network measures reveal how easy accessible information is in the reading network and how information is integrated in the reading network. They thus might provide another and complementary perspective on differences in reading processes (Ma et al., 2023) and global text processing.

The second research question addressed how the microscopic network measures are related to reader and text characteristics. Regarding reader characteristics, the results demonstrated that strength and betweenness centrality are related to WMC. Strength reflects how much a sentence has been visited (Siew, 2020), and thus this finding indicates that readers with high WMC pay more frequent visits to different sentences during reading than readers with low WMC. Since readers with higher WMC have more processing capacity, one could assume that they would be more efficient in integrating information across the text and visit sentences fewer, not more, times (Daneman & Carpenter, 1980; Daneman & Merikle, 1996). However, the present finding suggests that high WMC is positively related to strength, possibly reflecting more strategic reading behavior. This is corroborated by the negative relation between betweenness, which identifies linking sentences that bridge information within the text (Siew, 2020) and WMC. This finding might be explained by the fact that readers with high WMC strategically look back to the more important areas without connecting all sentences to each other (Kaakinen et al., 2002; Moss et al., 2023). Previous research demonstrated that readers with high WMC increased attention to important text areas (Moss et al., 2023) and are more strategic in allocating attention during reading (Kaakinen et al., 2002). WMC is also needed to make strategic decision during reading such as rereading to build a high-quality mental representation (Daneman & Merikle, 1996).

Concerning text characteristics, the current study also examined the relation between network measures and topic structure of the text. Strength, harmonic centrality, and local clustering coefficient were related to the topic structure of the text. The results showed that topic-medial sentences had higher strength compared to topic-initial and topic-final sentences. Topic-medial sentences contain the main body of the text, and thus it is plausible that these sentences also attract most of the transitions. Previous research also demonstrated that the main body of the text, elaborating on the main points, attracts more attention and integration (Goldman & Saul, 1990). Topic-final sentences demonstrated a higher harmonic centrality than topic-medial sentences, and thus readers transition less often from topic-ending sentences to other parts of the text compared to medial sentences. The local clustering coefficient of topic-final sentences is lower compared to the other sentences, indicating that less rereading is taking place around the topic-final sentence. Previous research has demonstrated that topic-ending sentences represent an important transition in the topic structure (Hyönä et al., 2002), and paying extra attention to these sentences is useful for a high-quality mental representation (Hyönä & Lorch, 2004; Hyönä et al., 2002). In addition, our research demonstrates that topic-ending sentences serve less as a starting point for integration, as shown by a lower betweenness centrality and local clustering coefficient. These conclusions do not exclude each other, as longer reading times do not correlate with betweenness centrality and local clustering coefficient.

Limitations and Future Research

Although the findings from this study provide more insight into how microscopic network measures are related to reader and text characteristics, there are some limitations that call for future research. A first limitation is that we did not include a performance measure on reading comprehension, and we were not able to link differences in reading performance to differences in microscopic network measures. For future research, we suggest to assess reading comprehension so that the relationship between microscopic network measures and reading performance can be examined as in the study of Ma et al. (2023). By doing so, the network metrics can be further validated by examining to what extent they correlate with the quality of the readers' mental representation.

A second limitation is that readers received a quite general learning task to read the texts (study as you would prepare for an exam). For future research it would be interesting to contrast different reading purposes in experimental designs and to examine to what extent this is reflected in different reading networks. Since there is a clear relationship between task demands and eye-tracking measures (Kaakinen & Hyönä, 2005, 2008, 2010), it would be interesting to examine how task demands are also reflected in these network measures. It has been shown that readers focus more on relevant information compared to irrelevant information in relation to the task they received and one might expect that microscopic network measures between relevant and irrelevant information might differ as well.

Another related suggestion for future research is to examine how microscopic network measures can reveal differences in text genres or other structural elements in texts with experimental designs. Ma et al. (2023) described that reading networks of expository texts will usually be more centralized than reading networks of narrative texts. However, empirical research is lacking so far on whether these network measures would be able to capture differences in text genres. Additionally, future studies could focus on whether microscopic network measures would also be able to capture differences between central versus peripheral ideas in the text. Central ideas are important to understand the overall meaning of the text better and peripheral ideas are less crucial to understand the text (Yeari et al., 2016). Another line of research could look at differences in network measures between consistent versus inconsistent information in the text (e.g., Hyönä et al., 2003). Inconsistencies interfere with updating the situation model, because inconsistent information contradicts earlier information and therefore creating a consistent integrated model is hampered.

For future research, it would also be interesting to further examine the network visualization of the readers' attention allocation during information intake while reading. Currently, there is research on gaze displays and how teachers interpret these displays (Knoop-van Campen et al., 2021). In these studies, gaze plots are used that are displayed on the reading material. However, it would be interesting to add the network visualization as an addition as it provides unique information on the integration processes which is less visible from a gaze plot.

Conclusion

In summary, this study demonstrated that a social network approach offers a novel perspective on moment-to-moment integration processes during reading as measured with eye tracking. The network measures calculated at the sentence level could be linked to individual differences and text structure. We believe that this social network approach is especially interesting when examining integration processes between more AOIs and could be further explored in other domains.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10648-025-10000-y.

Funding This research was supported by a grant from the FWO (G046219N).

Declarations

Conflict of Interest The authors declare no competing interests.

References

Alemdag, E., & Cagiltay, K. (2018). A systematic review of eye tracking research on multimedia learning. *Computers & Education, 125*, 413–428. https://doi.org/10.1016/j.compedu.2018.06.023

Ariasi, N., & Mason, L. (2011). Uncovering the effect of text structure in learning from a science text: An eye-tracking study. *Instructional Science, 39*(5), 581–601. https://doi.org/10.1007/s11251-010-9142-5

Ariasi, N., Hyönä, J., Kaakinen, J. K., & Mason, L. (2017). An eye-movement analysis of the refutation effect in reading science text. *Journal of Computer Assisted Learning*. https://doi.org/10.1111/jcal.12151

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press.

Barrat, A., Barthélemy, M., Pastor-Satorras, R., & Vespignani, A. (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences, 101*(11), 3747–3752. https://doi.org/10.1073/pnas.0400087101

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Borgatti, S. P. (2005). Centrality and network flow. *Social Networks, 27*(1), 55–71. https://doi.org/10.1016/j.socnet.2004.11.008

Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network analysis in the social sciences. *Science, 323*(5916), 892–895. https://doi.org/10.1126/science.1165821

Borgatti, S. P., Everett, M. G., & Johnson, J. C. (2013). *Analyzing social networks*. Sage.

Bormans, L. (2016). *The world book of hope*. Lannoo.

Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Maechler, M., & Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal, 9*(2), 378–400. https://doi.org/10.32614/RJ-2017-066

Castro, N., & Siew, C. S. Q. (2020). Contributions of modern network science to the cognitive sciences: Revisiting research spirals of representation and process. *Proceedings of the Royal Society A: Mathematical, physical and engineering sciences, 476*(2238), 20190825. https://doi.org/10.1098/rspa.2019.0825

Catrysse, L., Gijbels, D., & Donche, V. (2018). It is not only about the depth of processing: What if eye am not interested in the text? *Learning and Instruction, 58*, 284–294. https://doi.org/10.1016/j.learninstruc.2018.07.009

Clemente, G. P., & Grassi, R. (2018). Directed clustering in weighted networks: A new perspective. *Chaos, Solitons and Fractals, 107*, 26–38.

Clinton, V., & van den Broek, P. (2012). Interest, inferences, and learning from texts. *Learning and Individual Differences, 22*(6), 650–663. https://doi.org/10.1016/j.lindif.2012.07.004

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates, Publishers.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*(5), 769–786. https://doi.org/10.3758/BF03196772

Csárdi, G., Nepusz, T., Traag, V., Horvát, S., Zanini, F., Noom, D., & Müller, K. (2024). *igraph: Network analysis and visualization in R*. https://doi.org/10.5281/zenodo.7682609

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*(4), 450–466. https://doi.org/10.1016/S0022-5371(80)90312-6

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3*(4), 422–433. https://doi.org/10.3758/bf03214546

Fox, E. (2009). The role of reader characteristics in processing and learning from informational text. *Review of Educational Research, 79*(1), 197–261. https://doi.org/10.3102/0034654308324654

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks, 1*, 215–239. https://doi.org/10.1016/0378-8733(78)90021-7

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press.

Hyönä, J., & Lorch, R. F. (2004). Effects of topic headings on text processing: Evidence from adult readers' eye fixation patterns. *Learning and Instruction, 14*(2), 131–152. https://doi.org/10.1016/j.learninstruc.2004.01.001

Hyönä, J., & Nurminen, A. M. (2006). Do adult readers know how they read? Evidence from eye movement patterns and verbal reports. *British Journal of Psychology, 97*(1), 31–50. https://doi.org/10.1348/000712605X53678

Hyönä, J., Lorch, R. F., & Kaakinen, J. K. (2002). Individual differences in reading to summarize expository text: Evidence from eye fixation patterns. *Journal of Educational Psychology, 94*(1), 44–55. https://doi.org/10.1037/0022-0663.94.1.44

Hyönä, J., Lorch, R. F., & Rinck, M. (2003). Eye movement measures to study global text processing. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research*. Elsevier Science.

Jarodzka, H., & Brand-Gruwel, S. (2017). Tracking the reading eye: Towards a model of real-world reading. *Journal of Computer Assisted Learning, 33*(3), 193–201. https://doi.org/10.1111/jcal.12189

Kaakinen, J. K. (2017). On-line measures of text processing. In M. F. Schober, D. N. Rapp, & M. A. Britt (Eds.), *Routledge handbook of discourse processes* (2nd ed., pp. 125–130). Routledge.

Kaakinen, J. K., & Hyönä, J. (2005). Perspective effects on expository text comprehension: Evidence from think-aloud protocols, eyetracking, and recall. *Discourse Processes, 40*(3), 239–257. https://doi.org/10.1207/s15326950dp4003_4

Kaakinen, J. K., & Hyönä, J. (2007). Perspective effects in repeated reading: An eye movement study. *Memory & Cognition, 35*(6), 1323–1336.

Kaakinen, J. K., & Hyönä, J. (2008). Perspective-driven text comprehension. *Applied Cognitive Psychology, 22*, 319–334. https://doi.org/10.1002/acp

Kaakinen, J. K., & Hyönä, J. (2010). Task effects on eye movements during reading [research support, non-U.S. gov't]. *J Exp Psychol Learn Mem Cogn, 36*(6), 1561–1566. https://doi.org/10.1037/a0020693

Kaakinen, J. K., Hyönä, J., & Keenan, J. M. (2002). Perspective effects on online text processing. *Discourse Processes, 33*(2), 159–173. https://doi.org/10.1207/s15326950dp3302_03

Kendeou, P., & O'Brien, E. J. (2018). Reading comprehension theories: A view from top down. In M. F. Schober, D. N. Rapp, & M. A. Britt (Eds.), *The Routledge handbook of discourse processes*. Routledge.

Kendeou, P., & Trevors, G. (2012). Quality learning from texts we read. What does it take? In J. R. Kirby & M. J. Lawson (Eds.), *Enhancing the quality of learning. Dispositition, instruction, and learning processes.* (pp. 251–314). Cambridge university press.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*, 163–182. https://doi.org/10.1037/0033-295X.95.2.163

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

Knoop-van Campen, C. A. N., Kok, E., van Doornik, R., de Vries, P., Immink, M., Jarodzka, H., & van Gog, T. (2021). How teachers interpret displays of students' gaze in reading comprehension assignments. *Frontline Learning Research, 9*(4), 116–140. https://doi.org/10.14786/flr.v9i4.881

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lenth, R. V. (2022). *emmeans: Estimated marginal means, aka least-squares means*. https://CRAN.R-project.org/package=em-means.

Linderholm, T. (2002). Predictive inference generation as a function of working memory capacity and causal text constraints. *Discourse Processes, 34*, 259–280. https://doi.org/10.1207/S15326950DP3403_2

Lüdecke, D., Ben-Shachar, M., Patil, I., Waggoner, P., & Makowski, D. (2021). Performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software, 6*(60), 31–39. https://doi.org/10.21105/joss.03139

Ma, X., Liu, Y., Clariana, R., Gu, C., & Li, P. (2023). From eye movements to scanpath networks: A method for studying individual differences in expository text reading. *Behavior Research Methods, 55*(2), 730–750. https://doi.org/10.3758/s13428-022-01842-3

McNamara, D. S. (2011). Measuring deep, reflective comprehension and learning strategies: Challenges and successes. *Metacognition and Learning, 6*(2), 195–203. https://doi.org/10.1007/S11409-011-9082-8

McNamara, D. S., & Magliano, J. (2009). Toward a comprehensive model of comprehension. *In The psychology of Learning and Motivation, 51*, 297–384. https://doi.org/10.1016/S0079-7421(09)51009-2. Elsevier Academic Press.

McVay, J. C., & Kane, M. J. (2012). Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention. *Journal of Experimental Psychology. General, 141*(2), 302–320. https://doi.org/10.1037/a0025250

Mézière, D. C., Yu, L., McArthur, G., Reichle, E. D., & von der Malsburg, T. (2024). Scanpath regularity as an index of reading comprehension. *Scientific Studies of Reading, 28*(1), 79–100. https://doi.org/10.1080/10888438.2023.2232063

Mikkilä-Erdmann, M., Penttinen, M., Anto, E., & Olkinuora, E. (2008). Problems of constructing mental models during learning from science text. Eye tracking methodology meets conceptual change. In D. Ifenthaler, P. Pirnay-Dummer, & J. Michael Spector (Eds.), *Understanding models for learning and instruction: Essays in honor of Norbert M. Seel* (pp. 63–79). Routledge.

Moss, C., Kwabi, S., Ardoin, S. P., & Binder, K. S. (2023). Eye movements and reading comprehension performance: Examining the relationships among test format, working memory capacity and reading comprehension. *Reading and Writing*. https://doi.org/10.1007/s11145-023-10428-0

Olsen, A. (2012). *The Tobii I-VT fixation filter*. Tobii Pro.

Olsson, P. (2007). *Real-time and offline filters for eye tracking* KTH Royal Institute of Technology].

Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks, 32*(3), 245–251. https://doi.org/10.1016/j.socnet.2010.03.006

Pander Maat, H., Kraf, R., van den Bosch, A., Dekker, N., van Gompel, M., Kleijn, S., Sanders, T., & van der Sloot, K. (2014). T-Scan: A new tool for analyzing Dutch text. *Computational Linguistics in the Netherlands Journal, 4*, 53–74.

Penttinen, M., Anto, E., & Mikkilä-Erdmann, M. (2012). Conceptual change, text comprehension and eye movements during reading. *Research in Science Education, 43*(4), 1407–1434. https://doi.org/10.1007/s11165-012-9313-2

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of construcitvely responsive reading*. Erlbaum.

Rayner, K. (1998). Eye Movements in reading and information processing: 20 Years of Research. *Psychological Bulletin, 124*(3), 372–422.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology, 62*(8), 1457–1506. https://doi.org/10.1080/17470210902816461

Siew, C. S. Q., Wulff, D. U., Beckage, N. M., & Kenett, Y. N. (2019). Cognitive network science: A review of research on cognition through the lens of network representations, processes, and dynamics. *Complexity, 2019*, 2108423. https://doi.org/10.1155/2019/2108423

Siew, C. S. Q. (2020). Applications of network science to education research: Quantifying knowledge and the development of expertise through network analysis. *Education Sciences*, *10*(4), 101. https://www.mdpi.com/2227-7102/10/4/101

Singer, M., Andruslak, P., Reisdorf, P., & Black, N. L. (1992). Individual differences in bridging inference processes. *Memory & Cognition, 20*(5), 539–548. https://doi.org/10.3758/BF03199586

Unsworth, N., & McMillan, B. D. (2013). Mind wandering and reading comprehension: Examining the roles of working memory capacity, interest, motivation, and topic experience. *Journal of Experimental Psychology Learning, Memory, and Cognition, 39*(3), 832–842. https://doi.org/10.1037/a0029669

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods, 37*(3), 498–505. https://doi.org/10.3758/bf03192720

van den Broek, P. (2010). Using texts in science education: Cognitive processes and knowledge representation. *Science, 328*(5977), 453–456. https://doi.org/10.1126/science.1182594

van Gog, T., & Jarodzka, H. (2013). Eye tracking as a tool to study and enhance cognitive and metacognitve processes in computer-based learning environments. In R. Azevedo & V. A. W. M. M. Aleven (Eds.), *International handbook of metacognition and learning technologies*. Springer.

Verhavert, S., van Daal, T., & Catrysse, L. (2020). eyeRead [computer software]. https://cran.r-project.org/web/packages/eyeRead

Von der Malsburg, T., & Vasishth, S. (2011). What is the scanpath signature of syntactic reanalysis? *Journal of Memory and Language, 65*(2), 109–127. https://doi.org/10.1016/j.jml.2011.02.004

Yeari, M., Oudega, M., & van den Broek, P. (2016). The effect of highlighting on processing and memory of central and peripheral text information: Evidence from eye movements *Journal of Research in Reading* https://doi.org/10.1111/1467-9817.12072

Zhang, D., Hyönä, J., Cui, L., Zhu, Z., & Li, S. (2019). Effects of task instructions and topic signaling on text processing among adult readers with different reading styles: An eye-tracking study. *Learning and Instruction, 64*, 101246. https://doi.org/10.1016/j.learninstruc.2019.101246

Zhu, M., & Feng, G. (2015). An exploratory study using social network analysis to model eye movements in mathematics problem solving. LAK '15: Proceedings of the fifth international conference on learning analytics and knowledge

All the authors have approved the final article.