EasyChair Preprint

№ 6154

Predicting Cohesive Comprehension Based on Individual Differences and Genre Effects

Lauren Flynn, Joseph Magliano, Kathryn McCarthy, Danielle McNamara and Laura Allen

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

July 26, 2021

Predicting Cohesive Comprehension Based on Individual Differences and Genre Effects

Lauren E. Flynn[1]

Joseph P. Magliano[2]

Kathryn S. McCarthy[2]

Danielle S. McNamara[3]

Laura K. Allen[1]

[1]Department of Psychology, University of New Hampshire

[2]Department of Learning Sciences, Georgia State University

[3]Department of Psychology, Arizona State University

Author Note

The authors declare that there are no conflicts of interest with respect to this preprint. This research was supported in part by IES Grant R305A190063 as well as the Office of Naval Research (Grants: N00014-20-1-2627 and N00014-19-1-2424). Opinions, conclusions, or recommendations do not necessarily reflect the view of the Department of Education, IES, or the Office of Naval Research.

Correspondence should be addressed to Lauren at: lef1008@wildcats.unh.edu

Abstract

This study examined how individual differences and genre were related to the cohesion of readers' think-aloud protocols. Participants (n=119) were instructed to think-aloud while reading a history and science text and subsequently completed reading skill and working memory assessments. Results from correlations and mixed-effects models revealed that working memory was related to cohesion for history texts whereas reading skill was related to cohesion for both texts. These findings indicate that the interaction between genre and individual differences may be used to model coherence-building processes during reading.

Keywords: reading comprehension, individual differences, think-aloud methods

Introduction

Successful text comprehension occurs when a reader has constructed a *coherent* and meaningful mental representation of a text (McNamara & Magliano, 2009). The coherence of this representation arises when individuals not only process and understand the individual facts in the text but also generate inferences and elaborations using their own experiences or prior knowledge (Magliano et al., 1999). One way that researchers have identified and examined these *coherence-building* processes is through the collection of think-aloud protocols during reading. Think-aloud procedures typically involve readers being intermittently interrupted and prompted to report their thoughts as they come to mind. This allows for the collection of readers' thoughts during the reading process while minimizing the interference of probes by waiting until the end of sentences, as opposed to disrupting them mid-sentence (Pressley & Afflerbach, 1995).

Recently, studies have begun to examine the linguistic properties of individuals' think-alouds and how these may relate to successful comprehension processes. In particular, studies indicate that the *cohesion* of think-alouds is associated with reading and comprehension skills (Allen et al., 2016, 2015). Thus, the cohesion of think-alouds provides insight into readers' level of comprehension, with highly connected ideas about the text corresponding to deeper levels of comprehension. However, this work has yet to examine how these cohesive properties manifest in different types of text (i.e., across genres) nor has it examined how they relate to individual differences in skills relevant to comprehension processes. A number of individual differences have been linked to a reader's likelihood of successfully developing coherent representations during reading; thus, examinations of how these relate to think-aloud properties may shed light on how such individual differences manifest during the reading process. The current study

focuses on two such individual differences that have been particularly prevalent in the text comprehension literature: working memory and reading skill.

Prior research has found positive relations between working memory and comprehension, indicating that readers with lower performance on working memory capacity tasks may struggle to integrate texts into their mental representations due to weaker attention-control mechanisms (Daneman & Carpenter, 1980; Kane & McVay, 2012). Similarly, general reading skill has also been shown to relate to comprehension, with more skilled readers engaging in more coherence-building processes and vice versa (Klauda & Guthrie, 2008). However, less research has examined the specific mechanisms underlying the relations between these individual differences (working memory capacity and reading skills) and comprehension processes. The goal of this study is to examine *how* these individual differences (i.e., reading skill, working memory) manifest within the cohesive properties of online reading comprehension (think-alouds) across various genres.

Method

Undergraduate students (n= 119) read both a history (i.e., labor workers or Civil War) and science (i.e., evolution or erosion) text. Instructions asked that students read the texts for comprehension. Subjects were randomly assigned one text from each genre and read each line-by-line on their computers. Students were asked to type what they were currently thinking about at the end of each sentence. Think-alouds that were less than 100 words were removed. Students also completed assessments of reading skills (i.e., Gates-MacGinite Reading Test; MacGinitie et al., 2000) and working memory capacity (i.e., using the computerized version of the Operation Span task; OSPAN; Turner & Engle, 1989).

Linguistic Analyses

We used TAACO, a natural language processing (NLP) tool, to analyze readers' think-aloud protocols (Crossley et al., 2016). We analyzed texts for sentence-level, paragraph-level, and synonym cohesion (see Table 1 for index descriptions).

Table 1. Cohesion Analysis Indices

	Description
Adjacent Sentence Overlap	Number of lemma types that occur at least once in the next sentence
Adjacent 2-Sentence Overlap	Number of lemma types that occur at least once in the next two sentences
Adjacent Paragraph Overlap	Number of lemma types that occur at least once in the next paragraph
Adjacent 2-Paragraph Overlap	Number of lemma types that occur at least once in the next two paragraphs
Sentence Overlap of Noun Synonyms	Number of noun lemma types that occur at least once in the next sentence
Sentence Overlap of Verb Synonyms	Number of verb lemma types that occur at least once in the next sentence
Paragraph Overlap of Noun Synonyms	Number of noun lemma types that occur at least once in the next paragraph
Paragraph Overlap of Verb Synonyms	Number of verb lemma types that occur at least once in the next paragraph

Here, sentence-level cohesion is intended to be indicative of local connections that readers make whereas paragraph-level cohesion is indicative of readers' generation of more distal connections. Finally, synonym cohesion was examined to indicate whether readers were making semantic connections that were not reflected in explicit word overlap measures.

Data Analyses

Correlations were first analyzed between individual difference and cohesion variables at the aggregate level and then separately by text genre. Mixed effects models were then used to examine potential interactions between genre and individual differences on cohesion.

Results

We first examined correlations between individual differences and cohesion at three levels (sentence, paragraph, synonym; see Tables 2 and 3). Only one cohesion variable

demonstrated a significant correlation with working memory scores: adjacent paragraph verb synonym overlap (r=-.137, p<.05). Conversely, all but two of the cohesion variables was significantly related to reading skill. These correlations indicate that the connections observed in readers' think-aloud responses are more strongly related to their reading skill than their scores on the working memory capacity task.

Table 2. Correlations between Cohesion indices and Working Memory

Cohesion Index	History	Science	Total
Adjacent Sentence Overlap	-.246**	.032	-.110
Adjacent 2-Sentence Overlap	-.241**	.063	-.091
Adjacent Paragraph Overlap	-.248**	.048	-.107
Adjacent 2-Paragraph Overlap	-.230*	.068	-.086
Adjacent Sentence Noun Synonym Overlap	-.094	.016	-.034
Adjacent Sentence Verb Synonym Overlap	-.265**	.066	-.126
Adjacent Paragraph Noun Synonym Overlap	-.127	.015	-.047
Adjacent Paragraph Verb Synonym Overlap	-.281**	.085	-.137*
p <.01**; p <.05*			

We next examined whether these correlations differed by genre. We found significant, weak negative relations between cohesion and working memory for history texts, but not for science texts. Conversely, significant, very weak to weak positive correlations were observed between cohesion and reading skill for both the history and science texts, albeit with stronger relations within the history genre. These results suggest that texts may have subtle differential relations to coherence-building processes across different genres.

Table 3. Correlations between Cohesion indices and Reading Skill

Cohesion Index	History	Science	Total
Adjacent Sentence Overlap	.194*	.150	.170**
Adjacent 2-Sentence Overlap	.197*	.166	.179***
Adjacent Paragraph Overlap	.231*	.166	.197**
Adjacent 2-Paragraph Overlap	.246**	.169	.206**
Adjacent Sentence Noun Synonym Overlap	.209*	.202*	.202**
Adjacent Sentence Verb Synonym Overlap	.099	.118	.104
Adjacent Paragraph Noun Synonym Overlap	.222*	.215*	.216***
Adjacent Paragraph Verb Synonym Overlap	.111	.131	.133

p <.0001***; p<.01**; p <.05*

Finally, linear mixed-effects models were used to examine interactions between individual differences and genre in the cohesion of participants' think-alouds. We selected the variables with the highest correlation from each category (i.e., sentence-level, paragraph-level, synonym cohesion).

We examined models of working memory and genre. For all three models, there was a significant interaction between working memory and genre on think-aloud cohesion ($p < .01$), indicating that relations between working memory and cohesion were moderated by genre. Conversely, we found no significant interactions between reading skill and genre. These results suggest that reading skill was consistently related to think-aloud cohesion, regardless of the genre of the text. Overall, these results indicate that both individual differences and genre impact the think-aloud cohesion at the sentence-level, paragraph-level, and synonym-levels.

Discussion

The current study leveraged NLP techniques to automatically analyze the cohesion of readers' think-alouds across two different texts (one history and one science text). A think-aloud

procedure was used to collect readers' *online* coherence-building processes. Cohesion was then calculated for these think-alouds at three levels: sentence-level, paragraph-level, and synonym cohesion. Correlation analyses indicated that think-aloud cohesion was significantly correlated with reading skill for both history and science texts, indicating that skilled readers generated more explicit connections amongst the thoughts generated in their think-aloud responses. This is in line with prior work that has found significant relations between reading skill and think-aloud cohesion (Allen et al., 2015).

Conversely, working memory was only related to think-aloud cohesion for history texts and the relation was negative rather than positive. Thus, in certain contexts, individuals with higher working memory capacity may establish fewer explicit connections during reading. However, future research should more systematically examine these relations between working memory, genre, and coherence-building processes to develop a more thorough understanding of how such individual differences impact the coherence-building process.

Overall, this study suggests individual differences in reading skill and working memory partially manifest in the connections that readers generate while thinking-aloud. Importantly, reading skill demonstrated stronger and more stable relations to cohesion compared to working memory. Collectively, these results have both theoretical and practical implications for reading theory and practice. First, the results suggest that these individual differences may operate in different ways across varied contexts that may have differential demands on skill and capacity. Second, it suggests that reading skill may be more related to the cohesion of think-aloud responses than working memory. Given that reading skill is a malleable skill (unlike working memory), this suggests that difficulties associated with coherence-building may be more easily

remedied through strategy instruction. However, this remains an open question to be explored in future research studies.

References

Allen, L. K., Jacovina, M. E., & McNamara, D. S. (2016). Cohesive features of deep text comprehension processes. In J. Trueswell, A. Papafragou, D. Grodner, & D. Mirman (Eds.), *Proceedings of the 38th Annual Meeting of the Cognitive Science Society in Philadelphia, PA,* (pp. 2681-2686). Austin, TX: Cognitive Science Society.

Allen, L. K., Snow, E. L., & McNamara, D. S. (2015). Are you reading my mind? Modeling students' reading comprehension skills with Natural Language Processing techniques. In J. Baron, G. Lynch, N. Maziarz, P. Blikstein, A. Merceron, & G. Siemens (Eds.), *Proceedings of the 5th International Learning Analytics & Knowledge Conference (LAK'15),* (pp. 246-254). Poughkeepsie, NY: ACM.

Crossley, S. A., Kyle, K., & McNamara, D. S. (2016). The tool for the automatic analysis of text cohesion (TAACO): Automatic assessment of local, global, and text cohesion. *Behavior Research Methods*, *48*(4), 1227–1237. https://doi.org/10.3758/s13428-015-0651-7

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

Kane, M. J., & McVay, J. C. (2012). What mind wandering reveals about executive-control abilities and failures. *Current Directions in Psychological Science*, *21*(5), 348–354. APA PsycInfo. https://doi.org/10.1177/0963721412454875

Klauda, S. L., & Guthrie, J. T. (2008). Relationships of three components of reading fluency to reading comprehension. *Journal of Educational Psychology*, *100*(2), 310–321. https://doi.org/10.1037/0022-0663.100.2.310

MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Dreyer, L. G. (2000). *Gates-MacGinitie Reading Test (4th ed.)*. The Riverside Publishing Company.

Magliano, J. P., Zwaan, R. A., & Graesser, A. (1999). The role of situational continuity in narrative understanding. In *The construction of mental representations during reading* (pp. 219–245). Lawrence Erlbaum Associates Publishers.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*(2), 127–154. https://doi.org/10.1016/0749-596X(89)90040-5