Optional ERIC Coversheet — Only for Use with U.S. Department of Education Grantee Submissions

This coversheet should be completed by grantees and added to the PDF of your submission if the information required in this form **is not included on the PDF to be submitted**.

INSTRUCTIONS

- Before beginning submission process, download this PDF coversheet if you will need to provide information not on the PDF.
- Fill in all fields—information in this form **must match** the information on the submitted PDF and add missing information.
- Attach completed coversheet to the PDF you will upload to ERIC [use Adobe Acrobat or other program to combine PDF files]—do not upload the coversheet as a separate document.
- Begin completing submission form at https://eric.ed.gov/submit/ and upload the full-text PDF with attached coversheet when indicated. Your full-text PDF will display in ERIC after the 12-month embargo period.

GRANTEE SUBMISSION REQUIRED FIELDS

Title of article, paper, or other content

Anchoring your bridge: the importance of paraphrasing to inference making in self-explanations

All author name(s) and affiliations on PDF. If more than 6 names, ERIC will complete the list from the submitted PDF.

Last Name, First Name	Academic/Organizational Affiliation	ORCID ID
McNamara, Danielle S.	Department of Psychology, Arizona State Univ	
Newton, Natalie	Department of Psychology, Arizona State Univ	
Christhilf, Katerina	Department of Psychology, Arizona State Univ	
McCarthy, Kathryn S.	Department of Learning Sciences, Georgia Sta	
Magliano, Joseph P.	Department of Learning Sciences, Georgia Sta	
Allen, Laura K.	Department of Educational Psychology, Univer	

Publication/Completion Date—(if *In Press,* enter year accepted or completed) 06 Jul 2023

Check type of content being submitted and complete one of the following in the box below:

- ☑ If article: Name of journal, volume, and issue number if available
- ☐ If paper: Name of conference, date of conference, and place of conference
- ☐ If book chapter: Title of book, page range, publisher name and location
- ☐ If book: Publisher name and location
- ☐ If dissertation: Name of institution, type of degree, and department granting degree

Discourse Processes, Vol 60, Issue 4-5

DOI or URL to published work (if available) https://doi.org/10.1080/0163853X.2023.2225757

Acknowledgement of Funding— Grantees should check with their grant officer for the preferred wording to acknowledge funding. If the grant officer does not have a preference, grantees can use this suggested wording (adjust wording if multiple grants are to be acknowledged). Fill in Department of Education funding office, grant number, and name of grant recipient institution or organization.

"This work was supported by U.S. Department of Education **[Office name]** National Center for Educat through **[Grant number]** R305A1900 to **Institution]** Arizona State U. The opinions expressed are those of the authors and do not represent views of the **[Office name]** National Center for Educa or the U.S. Department of Education.

Routledge
Taylor & Francis Group

Discourse Processes

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hdsp20

Anchoring your bridge: the importance of paraphrasing to inference making in self-explanations

Danielle S. McNamara, Natalie Newton, Katerina Christhilf, Kathryn S. McCarthy, Joseph P. Magliano & Laura K. Allen

To cite this article: Danielle S. McNamara, Natalie Newton, Katerina Christhilf, Kathryn S. McCarthy, Joseph P. Magliano & Laura K. Allen (2023) Anchoring your bridge: the importance of paraphrasing to inference making in self-explanations, Discourse Processes, 60:4-5, 337-362, DOI: 10.1080/0163853X.2023.2225757

To link to this article: https://doi.org/10.1080/0163853X.2023.2225757

Published online: 06 Jul 2023.

DISCOURSE PROCESSES
2023, VOL. 60, NOS. 4–5, 337–362
https://doi.org/10.1080/0163853X.2023.2225757

Anchoring your bridge: the importance of paraphrasing to inference making in self-explanations

Danielle S. McNamara[a], Natalie Newton[a], Katerina Christhilf[a], Kathryn S. McCarthy[b], Joseph P. Magliano[b], and Laura K. Allen[c]

[a]Department of Psychology, Arizona State University; [b]Department of Learning Sciences, Georgia State University; [c]Department of Educational Psychology, University of Minnesota

ABSTRACT

Analyzing constructed responses, such as think-alouds or self-explanations, can reveal valuable information about readers' comprehension strategies. The current study expands on the extant work by (1) investigating combinations and patterns of comprehension strategies that readers use and (2) examining the extent to which these patterns relate to individual differences and comprehension outcomes. We leveraged archival data from three datasets (n = 472) to examine how comprehension strategy use varied across datasets, texts, and populations (high school, undergraduate). Students' self-explanations were coded for strategy use and then further analyzed in terms of combinations and patterns of strategies. Our analyses revealed that almost all readers primarily engaged in paraphrasing and/or the combination of paraphrasing and bridging, with few instances of elaboration. Further, the combination of paraphrasing and bridging was the best predictor of performance on a comprehension test. In terms of patterns, switching between strategies was not correlated to reading comprehension and was negatively correlated with the combination of paraphrasing and bridging. Understanding which strategy combinations and patterns are optimal can be used to inform adaptive instruction and feedback that can aid in more individualized support for readers.

Introduction

Comprehension of a written text emerges through converting words and sentences into ideas, connecting ideas across the text, and embedding them within the context of one's own prior knowledge. Key to this process is the generation of inferences to link information from across the text and to integrate information from prior knowledge: Active meaning construction and coherence-building are more likely to emerge when readers connect text ideas by generating knowledge-based and text-based inferences (Graesser et al., 1994; Kintsch, 1988, 1998; Palincsar & Brown, 1984; Scardamalia & Bereiter, 1992). Deep comprehension emerges when readers make inferences to develop a *coherent* mental model of the text (K. K. Millis et al., 2018; Van den Broek & Helder, 2017).

Readers are assumed to coordinate the use of different processes to support coherence-building across different parts of the text (e.g., McNamara & Magliano, 2009a; Rapp & van Den Broek, 2005; Van den Broek et al., 2015, 1999). Our objective in this study was to contribute to a better understanding of coherence-building and, in particular, how comprehension strategies are coordinated to support coherence-building. We leveraged analyses of constructed responses, specifically self-explanations, that were collected in the context of three prior studies with high school and college students. Constructed responses refer to the typed, written, or verbally produced responses that readers generate while reading a text, including thinking aloud and self-explanations.[1] Self-

CONTACT Danielle S. McNamara dsmcnamara1@gmail.com Department of Psychology, Arizona State University, 950 S. McAllister, Tempe, AZ 85287-1104

explanation in the context of comprehension tasks is the process of explaining text to oneself either orally or in writing, typically by typing the response. Analyzing comprehension strategies via constructed responses provides valuable insights into the cognitive processes underlying comprehension (McNamara, 2009; Rapp et al., 2007). Specifically, analyses of the self-explanations collected in our targeted studies allow us to examine students' use of individual strategies and combinations of strategies, how they switch between strategies from one response to the next, and how strategy use relates to text comprehension and individual differences in literacy and prior knowledge.

In the following introductory sections, we further elaborate on the literature that sets the stage for the current study. We discuss the distinction between comprehension processes and strategies, describe various types of processing strategies, and provide examples of how strategies can be used in combination.

Comprehension processes versus strategies

The discourse processing literature generally distinguishes between comprehension processes and comprehension strategies. Comprehension processes comprise the broad array of underlying mental processes that are necessary to understand text and discourse (e.g., lexical decoding, sentence parsing, backward semantic mapping). The success of comprehension processes varies based on individual differences among readers, which can be specific to comprehension skills (e.g., basic reading skills and vocabulary; Farley & Elmore, 1992; Nagy, 2007), domain-general knowledge and skills (e.g., working memory, metacognitive awareness, need for cognition, world knowledge; Dai & Wang, 2007; Just & Carpenter, 1992; Talwar et al., 2018), and domain-specific knowledge and skills (Alexander et al., 1994; McCarthy & McNamara, 2021).

Comprehension strategies refer to various types of reading behaviors in which readers may engage, which in turn evoke different comprehension processes. There are two broad classes of comprehension strategies: task-level strategies and processing strategies. Task-level strategies are deliberate activities designed to support and enhance understanding and learning from text (McNamara, 2007). There are a multitude of task-level comprehension strategies, including prereading, notetaking, summarizing, self-explanation, generating and answering questions, discussions, and composing argumentative essays. An underlying key component to most effective comprehension strategies is the degree to which they support inference generation before, during, and after reading (Van den Broek et al., 2015).

Processing strategies refers to the online, more moment-to-moment strategies that readers employ during reading, such as *paraphrasing*, generating *bridging* inferences to connect ideas in the text, and *elaborative* inferences that expand on the text content using prior knowledge. Some inferences are automatically generated based on the activation of concepts in memory and prior knowledge (Kintsch, 1988), but inferencing can also be under the strategic control of the reader (Graesser et al., 1994; Magliano et al., 1999a; Van den Broek et al., 2001, 1999). Likewise, various processing strategies can emerge naturally within the context of reading: Skilled readers are more likely to use inference strategies, such as bridging and text-relevant elaborations (Magliano et al., 2020; McCrudden et al., 2021), whereas less skilled readers are more likely to rely on less effective strategies, such as paraphrasing and tangential elaborations (Carlson et al., 2014; Long et al., 1999).

Various methods can be used to assess the extent to which readers engage in different processing strategies, including reading time, lexical decision tasks, eye-tracking, and constructed responses (McCarthy et al., 2018). Evidence for inferences to connect ideas is generally interpreted as a signal that the reader is engaging in more or less strategic processing of the text. Response latencies, reading time, and eye-tracking methodologies are arguably less invasive and more widely used to study bridging or elaborative inference processes (Magliano & Graesser, 1991; Potts et al., 1988). While these methodologies may be informative of comprehension processing in some study designs, constructed responses are valuable for understanding comprehension because they provide a record of the consciously available thoughts that are the products of strategic processes (Ericsson & Simon, 1993; Trabasso & Magliano, 1996). Such responses often contain evidence of multiple strategies (e.g., Trabasso & Magliano, 1996) and, as such, afford an exploration of how strategies occur and co-occur, which is not readily afforded by processing time or eye-tracking data alone (Magliano & Graesser, 1991).

Constructed response instructions can also serve to specifically prompt task-level strategies that tend to bias particular processing strategies. Asking readers to explain to themselves as they read, or prompting *self-explanation*, reveals students' propensity to use comprehension strategies that support inference generation. In turn, self-explanation instructions can prompt students to leverage effective comprehension strategies more characteristic of skilled readers by inducing a focus on meaning making and an understanding of causal relations while reading (Coté & Goldman, 1999; Koslowski & Masnick, 2002; Lombrozo, 2006). When students produce high quality self-explanations (either spontaneously or following an instructional prompt), they understand more from learning materials and construct better mental models of the content (Chi & VanLehn, 1991; Chi et al., 1989, 1994; Magliano et al., 1999a; Trabasso & Magliano, 1996; VanLehn et al., 1992). By contrast, instructions to generate surface-level responses to a text (e.g., paraphrasing) do not necessarily prompt students to make connections that are characteristic of skilled readers (Coté & Goldman, 1999; Millis et al., 2007).

Processing strategies: the importance of inferencing

The inferences involved in comprehension can be characterized as two broad classes: *bridging* and *elaborative* (Seifert et al., 1985; Singer, 1988). Bridging inferences involve establishing connections between explicit discourse constituents. These can involve lower-level features such as the resolution of anaphora or higher-level situational relations, such as causal, motivational, temporal, spatial, logical, and argumentative relations. Elaborative inferences, on the other hand, involve readers drawing on knowledge that goes beyond the discourse context and relies on prior knowledge (Graesser & Clark, 1985; Graesser et al., 1994; McNamara & Magliano, 2009b; O'Brien et al., 1988; Singer, 1988).

There is considerable evidence that bridging and elaborative inferences are both critical to comprehension (Magliano & Millis, 2003; Magliano et al., 2011; Millis et al., 2006; Singer et al., 1992; Whitney et al., 1991). Readers who tend to generate inferences while reading demonstrate better memory of the text than those with a greater focus on local information by rereading, repeating the text, or paraphrasing each sentence (Magliano et al., 1999a). In turn, readers' use of comprehension strategies is fundamentally related to comprehension skill (e.g., Magliano et al., 2011). Readers who tend to bridge to previously read text content perform better on separate comprehension skill assessments compared to those who tend to paraphrase the content (Carlson et al., 2014; Magliano & Millis, 2003; McMaster et al., 2012).

Notably, paraphrasing has often been considered as a signature of surface level and less skilled reading (Carlson et al., 2014; Magliano & Millis, 2003; cf. McNamara et al., 2009). However, paraphrasing is nonetheless an important strategy for younger readers (Karlsson et al., 2018) and readers who are struggling to understand the content of the text (McNamara, 2004). Readers who generate higher quality paraphrases better understand text (Haynes & Fillmer, 1984). Likewise, providing instruction on how to paraphrase enhances young, developing readers' comprehension skills (cf. McNamara et al., 2009).

Rephrasing the text into words that are more familiar and accessible affords the reader a better understanding of the explicit meaning of the text, albeit at the local level (Kintsch & van Dijk, 1978). Even for more skilled readers, generating a bridging or elaborative inference without generating a foundational paraphrase is likely to result in a less coherent mental representation of the content (McNamara, 2009). As such, while *solely* paraphrasing is considered to be less strategic, there are few studies that have examined paraphrasing in combination with other strategies.

Combinations of processing strategies

Readers may differ not only in the types of strategies they use, but also in their patterns of strategy use across the text and their tendency to switch between strategies. For example, when readers self-explain, they deploy a variety of strategies in combination (McNamara, 2004; McNamara & Magliano, 2009a). Consider Table 1, which illustrates constructed responses, specifically self-explanations produced while reading a text on cell division (see the OSF repository,[2] Appendix A for the full texts).

Example 1 shows that the reader paraphrased content in the target sentences (i.e., rephrasing content for the sentence that was just read) and made a bridging inference to connect the new information to content conveyed in a prior sentence. In Example 2, the reader elaborated on the conditions necessary for growth of the cytoplasm. The reader paraphrased content from the sentence and elaborated based on world knowledge regarding conditions that lead to the growth of biological systems. In Example 3, the reader bridges to content conveyed in the prior text (e.g., there are additional steps), elaborates based on prior knowledge (e.g., Humans have organs), and generates a paraphrase. These examples demonstrate that readers are not merely paraphrasing *or* bridging *or* elaborating but, rather, are likely to combine strategies. All three examples show that coherence-building can involve an *anchoring* of bridging and elaboration to the content that was just read (i.e., the target sentences where the protocols were produced), which requires a coordination of strategies when producing the constructed responses. An anchor prevents a floating craft from drifting due to wind or current. In the same way, paraphrases can serve to connect an inference to the reader's mental model. As illustrated in Table 2, connections to the current sentences (i.e., paraphrases) can provide anchors at the beginning, middle, or end of a constructed response.

There has been little work examining readers' combinations and patterns of strategies across a reading task (cf. Cromley & Wills, 2014; Goldman & Saul, 1990; Karlsson et al., 2018; Seipel et al.,

Table 1. Example constructed responses that illustrate a coordination of coherence-building strategies.

Example	Target Sentence	Constructed Response
1	The place where they are attached is called the centromere.	The 92 chromatids in a human cell *are attached at one specific spot called the centromere*. This appears as 46 pairs of chromatids.
2	Therefore, before cytokinesis, there is growth in the size of the cytoplasm and in the number of its organelles.	Since *cytokinesis* depends on certain variables, *there is a growth in size of the cytoplasm and number of organelles* to help the cytokinesis. It needs energy and other variables to continue, just as anything in an organism does.
3	Each daughter cell will have about half of the organelles from the cytoplasm of the parent cell	Like humans cells have organs that allow the cells to complete their functionality. *Each cell is given half of their needed organs from the parent cell.* There is still another step to be done before each cell is complete.These cells are floating around in the goo that is the cytoplasm.

Each constructed response includes information from the target sentence (underlined). Each response also expands beyond the target sentence, through including prior knowledge or connecting to earlier parts of the text. These example responses therefore demonstrate anchored coherence-building, meaning coordinated use of paraphrasing and one or more additional strategies.

Table 2. Examples of paraphrasing and bridging for the cell division text.

Text*	Self-Explanation**
In eukaryotic cells there are two distinct but overlapping stages of cell division. **In the first stage, mitosis, one complete set of chromosomes goes to each daughter cell.**	In the first stage *of cell division in eukaryotic cells*, each daughter cell which is a different cell gets one complete set of chromosomes, this process is called mitosis
	With eukaryotic cells, there is a division, and the first part is mitosis, where one set of chromosomes goes to each cell.
	The first stage *of cell division* is called mitosis. In this stage, a complete set of chromosomes goes to each daughter cell. *This is one of two distinct but overlapping stages of cell division.*

The target sentence is shown in boldface. Underlined text indicates a paraphrase attempt; *italicized* text indicates a bridging attempt.

2017) and limited (or no) work investigating how readers combine strategies within a constructed response. For example, some studies have adopted a relatively holistic approach by focusing on the quality of the response or coding a response in terms of the predominant strategy used (e.g., Coté et al., 1998; Van den Broek et al., 2001). Some studies have adopted a more fine-grained approach by parsing each response clause into idea units and then categorizing each idea unit as a strategy (e.g., Dahl et al., 2021; Goldman et al., 2012; Karlsson et al., 2018; McNamara, 2004; Park et al., 2020; Trabasso & Magliano, 1996). Idea units are typically defined as verb clauses (e.g., Trabasso & Magliano, 1996), and at this grain size the strategy categories are treated as mutually exclusive (i.e., a verb clause can reflect only one strategy). While such an approach allows for consideration of multiple strategies within a single response, analyses of these codes are still typically conducted at the aggregate. Neither approach is well suited to exploring how strategies are combined and have tended to focus on characterizing the reader in terms of their propensity to engage in key strategies (Anmarkrud et al., 2014; Best et al., 2004; Karlsson et al., 2018; Wang, 2016), or they have reported patterns of strategy use and make arguments about the nature of comprehension (e.g., Trabasso & Magliano, 1996). As such, strategy use is typically analyzed at the aggregate for each participant to provide mean scores for individual strategies, such as how much of the reader's processing was dedicated to paraphrasing or elaborating, or the extent to which readers can be classified or profiled as a function of their tendency to engage in strategies (e.g., paraphrasers, elaborators, or lateral connectors; Carlson et al., 2014; Karlsson et al., 2018; McMaster et al., 2012; Rapp et al., 2007). Such approaches are useful but prevent analysis of how processing strategies might co-occur and how readers dynamically apply different strategies over the course of a text (cf., Coté et al., 1998; Cromley & Wills, 2014; Seipel et al., 2017).

Thus, while studies have shown that bridging, elaboration, and paraphrasing are correlated with comprehension outcomes (e.g., Magliano & Millis, 2003; Magliano et al., 2011, 2020; Millis et al., 2006), few studies have explored how strategies *combine* to support comprehension. This constitutes a considerable gap in the literature. While there is reason to suspect that flexible strategy use is beneficial for comprehension (Magliano et al., 1999a), to date, little evidence supports that supposition. On one hand, skilled readers may be more sensitive to differences within a text and therefore more likely to adjust or switch their strategy use to match the text. On the other hand, constantly shifting between strategies could be inefficient, indicating participants are not being thoughtful in how they explain a text. Thus, additional exploration into patterns of strategy use is warranted. We address this gap in the current work by exploring combinations of strategies within constructed responses while reading.

Current study

This study uses constructed responses to examine readers' use of comprehension strategies while self-explaining science texts. Here we focus on online comprehension processes during reading, and in particular the extent to which the reader is focusing on the current sentence and generating inferences to connect ideas. Our aim was to characterize comprehension behaviors exhibited by high school and college readers with particular focus on understanding readers' combination and coordination of strategies as they move through a text. This exploratory analysis of archival data comprises three datasets (see Tables 3 and 4), including 242 high school students (Dataset 1: McCarthy et al., 2018), 77 undergraduate college students (Dataset 2: Allen et al., 2017), and 153 undergraduate students (Dataset 3: Creer et al., 2020). Because comprehension, reading skills, and prior knowledge were assessed in the three studies, we were able to examine how strategy use relates to comprehension as well as individual differences in literacy and prior knowledge.

We examine students' strategy use at the response level by coding the presence of the three focal comprehension strategies (paraphrasing, bridging, elaboration). This approach affords the analysis of the use of multiple strategies within the same response as well as whether students switch between strategies (or strategy combinations) from one response to the next. We used a simple gauge of strategy flexibility by examining how often readers switched comprehension strategies between each self-explanation response. To explore readers' flexibility beyond overall amounts of switching between

Table 3. Description of participants and materials per dataset.

Variable	Dataset 1	Dataset 2	Dataset 3
Participants	High school students	Undergraduate students	Undergraduate students
Number	242	77	153
Age or Year in School	$^{M}Age = 17.59$ $^{SD}Age = 1.39$	52 Freshman 16 Sophomore 8 Junior 1 Senior	$^{M}Age = 21.73$ $^{SD}Age = 5.16$
Gender	150 Female 89 Male 3 NA	36 Female 41 Male	112 Female 39 Male 2 NA
Race	117 Caucasian 26 African American 55 Hispanic 19 Asian 22 Other 3 NA	34 Caucasian 1 African American 10 Hispanic 28 Asian 3 Other 41 NA	122 Caucasian 8 African American 7 Hispanic 8 Asian 8 Other
Text(s)	Heart Disease Red Blood Cells	Cell Division	Cell Division
Comprehension Test	Open-ended	Open-ended	Open-ended
Text-Based Questions	4	6	6
Bridging Inference Questions	4	6	6
Cohen's Kappa	.85	.91	.79
Individual Difference Measures	Prior Knowledge: 30 multiple-choice questions on science, history, and literature Gates-MacGinitie Reading Comprehension (Form S, Level 10/12)	Prior Knowledge: 30 multiple-choice biology questions	Prior Knowledge: 30 multiple-choice questions on science, history, and literature Prior Knowledge: 29 multiple-choice biology questions Gates-MacGinitie Vocabulary (Form S, Level 10/12)

Additional information about the studies is available from the following sources: Dataset 1 (McCarthy et al., 2018); Dataset 2 (Allen et al., 2017); Dataset 3 (Creer et al., 2020). The texts, comprehension tests and prior knowledge measures are provided in the OSF repository.

[a]Dataset 1 has a larger number of participants than reported in McCarthy et al. (2018), which excluded some participants due to attrition and missing data that are not relevant to the current paper.

Table 4. Description and features of each text.

Measure	Heart Disease	Red Blood Cells	Cell Division
Words	311	282	649
Sentences	21	20	48
Paragraphs	4	4	5
Target Sentences	9	9	16
Flesch Kincaid Grade Level	7.79	8.99	9.25
Coh-Metrix Narrativity Percentile	34	20	12
Coh-Metrix Syntactic Simplicity Percentile	74	94	82
Coh-Metrix Word Concreteness	86	88	33
Percentile Coh-Metrix Referential Cohesion Percentile	74	87	84
Coh-Metrix Deep Cohesion Percentile	98	98	17

strategies, we also leveraged Markov chains to demonstrate the specific types of strategy *transitions* that occur. Markov chains are used to demonstrate the probabilities of transitions from one process or event to another. Hidden Markov models have been frequently used to model student learning through Bayesian Knowledge Tracing (Corbett & Anderson, 1995; Nguyen et al., 2020; Williamson & Kizilcec, 2021). However, they have not been widely used to study readers' strategy use and their relation to text comprehension (cf. Christhilf et al., 2022). In our case, the strategy used in one response may depend on the strategy used within a previous response. Hence, these probabilities are determined by the frequencies of transitions in the data, which are displayed using a visual diagram.

Readers' self-explanations were coded using a common rubric and team of raters (see OSF repository, Appendix B, for full coding rubric; McCarthy et al., 2021). Across research groups who have examined strategy use within constructed responses, teams have used similar but not identical rubrics and approaches. Thus, one limitation in extant work is that different scoring approaches have been employed across studies, resulting in natural variations across studies. The current study is thus relatively unique in being scored by the same team, under the same guidelines, and using the same rubric. Such an approach allows for more clarity around how processing potentially varies or is comparable across texts, readers, and studies.

Methods

Datasets

Methods and procedures

The archival data were collected from three datasets (Dataset 1: McCarthy et al., 2018; Dataset 2: Allen et al., 2017; Dataset 3: Creer et al., 2020). Table 3 describes the properties of the participants and the measures used for each dataset. In all three studies, participants were given instructions to self-explain target sentences while reading a relatively challenging science text. Each participant read and self-explained a single science text. Participants in Dataset 1 were randomly assigned to one of two science texts (heart disease or red blood cells), whereas participants in Datasets 2 and 3 all read cell division.

The self-explanation instructions were similar across the three target studies. For example, in Datasets 2 and 3, participants were instructed as follows: *You will now be asked to read a text. After you read, your understanding will be assessed. To help you with this task, we would like you to provide your self-explanations of the text while you read. Self-explanations can help to improve your own understanding of the text. After some sentences, which are bolded, there is a blank box. For each of these bolded sentences, please read the text, and then write an explanation in the space provided below the segment. Please note that there are no "right" or "wrong" self-explanations.* Participants were provided examples of self-explanations, but no explicit training (or feedback) on how to use comprehension strategies to improve the quality of the self-explanation. See the OSF repository, Appendix F, for the full instructional prompts for each dataset. At predetermined sentences, participants were prompted to self-explain. Participants self-explained 1 of every 2 to 3 sentences because prompting participants to explain every sentence results in a fatigue effect and induces a greater focus on the target sentences (e.g., paraphrasing) and a lesser focus on making connections across the sentences (Kurby et al., 2012).

Dataset 1 included high school students, whereas Datasets 2 and 3 included undergraduate students. As such, the target texts in Dataset 1 (heart disease, red blood cells) were shorter and less challenging than the cell division text used in Datasets 2 and 3. Dataset 1 included the self-explanations generated during the pretest of a larger training study (McCarthy et al., 2018) wherein the participants read one of two counterbalanced texts matched for linguistic difficulty, at pretest and the other text during the posttest. In the current study, we only examined the pretest data so that the context is comparable to Datasets 2 and 3, which did not include comprehension strategy training.[3]

Assessments

Each study included an individual difference measure to assess prior knowledge of biology, general prior knowledge, or both (see the OSF repository, Appendices D and E). Dataset 1 included the Gates MacGinitie Reading Comprehension test, and Dataset 3 included the Gates MacGinitie Vocabulary test (MacGinitie & MacGinitie, 1989).

All studies included a comprehension test on the text that the participants self-explained, which was presented after reading the text. See the OSF repository, Appendix C, for the complete set of comprehension test questions. These questions were open-ended and were scored for partial credit (0.5) for incomplete but correct information and for full credit (1) with complete, correct information. No credit (0) was awarded for answers that were inaccurate or irrelevant to the question. Reliability in coding the responses was strong for each of the three studies (Cohen's kappa: Dataset 1 = .85; Dataset 2 = .91; Dataset 3: .79). The comprehension tests included an equal number of textbase questions and bridging inference questions (i.e., Dataset 1 = 4 each; Dataset 2 = 6 each; Dataset 3 = 6 each). The answers to textbase questions can be found in a single sentence in the text. In contrast, bridging inference questions probe for deeper comprehension as they require the reader to connect information across two or more sentences in the text to derive the answer.

Texts

The students in the three studies read and self-explained relatively prototypical science texts that might be found in a biology textbook. Heart disease explained common heart problems and how they impact the heart's functions. Red blood cells described the properties and function of red blood cells, as well as what happens when they do not work effectively. Cell division described the process of mitosis in eukaryotic cells.

Table 4 provides descriptive indices and Coh-Metrix indices, which represent features related to text readability (McNamara et al., 2014). All three texts are informational texts and thus lower in narrativity (e.g., fewer events, pronouns, and adverbs) compared to narrative or history texts. Informational texts often include high levels of referential cohesion but are generally more challenging to understand because they include fewer familiar concepts and readers tend to have less prior knowledge about the concepts (McNamara et al., 2012). Heart disease and red blood cells ware relatively brief texts (i.e., ~300 words) and were comparable in terms of all of the indices. Cell division (see also, McNamara, 2001) was longer and more challenging than the other two texts. Specifically, cell division was lower in narrativity (and thus more informational) and used fewer concrete words (and thus was more abstract). Its lower deep cohesion score compared to the other two texts indicated that there are few connectives used to explicitly describe the relations between ideas.

Table 5. Questions driving the dimensions of interest for each strategy.

Dimension	Paraphrasing	Bridging	Elaboration
Presence of Strategy	How much of the target sentence is captured? *(none; some; most)*	To what extent is information from previous text present? *(none, little, some, much)*	To what extent does the reader bring in outside information to make sense of the text? *(none, some, much)*
Nature of Strategy Use	To what extent is information rephrased in terms of semantics and syntax?	Are connections made to ideas in the previous sentence (*local bridge*) or other previous sentences (*distal bridge*)?	Are connections made to ideas in the previous sentence (*local bridge*) or other previous sentences (*distal bridge*)?
Overall Quality	What is the overall *quality* of the response? *(Poor, Fair, Good, Great)*		

The full rubric is provided in the OSF repository, Appendix B (see also McCarthy et al., 2021).

Coding comprehension strategies

The self-explanations were scored by expert raters using the Self-Explanation Rubric. The full rubric can be found in the OSF Repository, Appendix B. Table 5 shows the three broad dimensions targeted in the rubric. Expert rater pairs attained inter-rater reliability with a weighted Cohen's kappa value of 0.71 or higher across categories and scored the self-explanations independently, with 10% to 20% overlap in responses to perform continuous reliability checks (McCarthy et al., 2021).

The rubric identified three focal strategies: paraphrasing, bridging, and elaboration. *Paraphrasing* is a process in which students reproduce content from the texts. *Bridging* is a process of establishing how the current sentence is related to prior discourse context and provides the primary basis for achieving coherence in a mental model. Bridging attempts included restating information from earlier in the text. Attempts at summarizing (without grounding the explanation in a specific sentence or with specific text details) were also considered bridging attempts. Finally, *elaboration* is a process that involves integrating information from prior knowledge with information provided by the text. Bridging and elaboration are *inferential processes* that are critical for mental model construction. Notably, the rubric did not require raters to categorize each protocol as *either* paraphrase, bridging, or elaboration but rather captured the extent to which these strategies were apparent in students' protocols.

The first dimension was the presence (1) or absence (0) of the strategy. The second captured more detail about the nature of the strategy by scoring qualities of the response. For example, in paraphrasing, one reader might produce a paraphrase that only changes one or two key contents words, whereas another might engage in greater transformation of the target sentence, such as changing the syntax. Thus, scores in this dimension reflect to what extent the reader used the strategy. Third, the responses were also given an overall quality score (0–3). We included this score to reflect the extent to which the reader produced explanations that explained the text segment at a global level, incorporated content from across the text, and/or incorporated prior knowledge.

Additional codes were included in the rubric to capture other common behaviors (e.g., comprehension monitoring, inaccurate statements, life events, direct coping or retyping). We report these in this study, but these behaviors are not the focus on the current study and are relatively infrequent.

Each self-explanation could be coded for the presence of multiple strategies affording the ability to examine the co-occurrence of strategies. The presence scores of the focal strategies were converted to assign each self-explanation to one of eight strategy combination categories: none (none of the strategies was used), paraphrase only, bridge only, elaboration only, paraphrase and bridging combined (para-bridge), paraphrase and elaboration combined, bridging and elaboration combined, and all (paraphrase, bridging, and elaboration combined). These strategy combinations only noted which strategies were present in a self-explanation, meaning that presence or absence of a strategy was converted to a dichotomous variable.

Preliminary exploration of the data revealed that the second dimension reflecting the nature of the strategy lacked variability –87% of self-explanations coded for bridging contained at least one full idea, 78% of responses coded for elaboration contained a full idea, and all responses coded for paraphrase contained a full idea. Thus, we did not include the second dimension in the analyses. For each participant, the frequency of each type of strategy combination was calculated in terms of the proportion to the total number of constructed responses (*strategy combination frequency/total constructed responses)*. This proportion is used in all correlation and regression analyses. Appendix A displays examples of each of these strategy combinations.

Transitions between strategies

We examined students' tendencies to transition between strategies in two ways. First, we computed a switching score to indicate how often a participant changed the strategy combination used from one self-explanation to the subsequent self-explanation. Each of the eight possible strategy combinations were considered a discrete type of strategy use. For example, if a participant produced an isolated paraphrase within their first self-explanation and then produced a paraphrase-bridge combination in their second self-explanation, this would be counted as a strategy switch. If the participant then returned to solely paraphrasing on their third response, this would count as a second switch. Participants were prompted to produce self-explanations at 9 target sentence locations in Dataset 1 and 16 target sentence locations in Datasets 2 and 3. Thus, participants in Dataset 1 could receive a switching score from 0 to 8, while those in Datasets 2 and 3 could receive a switching score between 0 and 15. In order to make comparisons across the datasets given the differences in number of target sentences, we computed proportion scores. For the switching score, proportions were calculated by *switches/(total targets – 1).*

We further examined participants' transitions between strategies by leveraging Markov chains (D'Mello & Graesser, 2010; Geyer, 1992). The frequency of switching from one strategy to another was counted for every possible strategy transition. Each frequency was then converted to a probability by dividing by the total number of transitions from each strategy type. For example, suppose there were 10 instances in which participants transitioned from paraphrasing to bridging and 100 instances overall in which participants transitioned from paraphrasing to a subsequent strategy. The probability of switching from paraphrasing to bridging would then be .10. To create a visual illustration of these probabilities, the overall frequencies of each type of strategy were visualized as nodes; the size of the nodes varied in proportion to the frequency (i.e., dividing the overall strategy frequency by pi, then taking the square root). The transitional probabilities were visualized as arrows connecting the nodes; higher probabilities were visualized as wider arrows. The frequency of each type of switch from one strategy to another was calculated. Each switch frequency was then divided by the sum of the overall number of switches from that category (e.g., dividing the frequency of switching from paraphrasing to bridging by the frequency of switching from paraphrasing to any strategy). This value was then used to determine the arrow widths.

Results

In the following, we first present preliminary analyses to describe student performance on the comprehension and individual difference measures and then describe the frequency of strategy used in the three studies. We then address the primary objectives of the current study. Our first objective was to characterize the self-explanation responses by examining the frequency of the focal strategies within each response as well as the propensity to switch between strategies across responses. We then examined the relations between strategy use, text comprehension, and individual difference measures to better understand how strategy is related to performance outcomes and prior abilities. To illustrate how students transitioned between strategies, we visualized these transitions using a Markov chain. Finally, we examined the extent to which strategy use accounts for variance in comprehension.

Preliminary analyses

Comprehension and individual differences

Tables 6–8 provide the means and correlations for the comprehension scores and individual differences corresponding to each of the three datasets. There are several commonalities. Performance on the textbase and bridging inference questions were moderately to strongly correlated ($r = 0.48$–0.76; overall $r = 0.60$), with consistently higher performance on the textbase questions ($M = 0.53$; $SD = 0.28$) than the bridging questions ($M = 0.39$; $SD = 0.25$). Thus, this result provides construct validity regarding the expected difference between textbase and bridging questions.

Comprehension performance was consistently correlated with individual difference measures, which provides evidence for the validity of the comprehension assessments. Specifically, in Dataset 1, reading skill and knowledge had equivalently high correlations with the comprehension measures ($r = 0.45$–0.53), with the strongest relations with overall comprehension. Likewise, prior knowledge in biology was strongly correlated with comprehension ($r = 0.58$–0.64) in Dataset 2. Prior domain

Table 6. Dataset 1 means, standard deviations, and correlations.

Variable	*M*	*SD*	1	2	3	4
1. Overall Comprehension	0.49	0.22				
2. Textbase Questions	0.58	0.27	.89** [.86, .91]			
3. Bridging Questions	0.38	0.23	.80** [.74, .84]	.49** [.38, .58]		
4. Reading skill	0.57	0.22	.53** [.43, .62]	.45** [.34, .54]	.49** [.39, .58]	
5. Prior General Knowledge	0.57	0.15	.53** [.43, .61]	.46** [.36, .56]	.46** [.35, .55]	.68** [.60, .74]

M and *SD* are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. The confidence interval is a plausible range of population correlations that could have caused the sample correlation (Cumming, 2014).
[a]Means and correlations for Heart Disease and Red Blood Cell were similar and were thus aggregated.
$^{*}p < .05$. $^{**}p < .01$.

Table 7. Dataset 2 means, standard deviations, and correlations.

Variable	*M*	*SD*	1	2	3
1. Overall Comprehension	0.35	0.26			
2. Textbase Questions	0.4	0.28	.94** [.91, .96]		
3. Bridging Questions	0.3	0.26	.93** [.90, .96]	.76** [.65, .84]	
4. Prior Biology Knowledge	0.36	0.17	.64** [.49, .76]	.63** [.47, .75]	.58** [.40, .71]

M and *SD* are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. The confidence interval is a plausible range of population correlations that could have caused the sample correlation (Cumming, 2014).
$^{*}p < .05$. $^{**}p < .01$.

Table 8. Dataset 3 means, standard deviations, and correlations.

Variable	*M*	*SD*	1	2	3	4	5
1. Overall Comprehension	0.49	0.23					
2. Textbase Questions	0.53	0.26	.93** [.91, .95]				
3. Bridging Questions	0.46	0.25	.93** [.90, .95]	.72** [.64, .79]			
4. Prior Knowledge General	0.58	0.18	.39** [.25, .52]	.41** [.27, .53]	.32** [.17, .46]		
5. Prior Biology Knowledge	0.42	0.15	.45** [.31, .56]	.45** [.32, .57]	.37** [.23, .50]	.59** [.47, .68]	
6. Vocabulary knowledge	0.68	0.19	.34** [.19, .48]	.37** [.22, .50]	.27** [.11, .41]	.66** [.56, .74]	.52** [.40, .63]

M and *SD* are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. The confidence interval is a plausible range of population correlations that could have caused the sample correlation (Cumming, 2014).
$^{*}p < .05$. $^{**}p < .01$.

knowledge related to biology similarly emerged as the strongest correlate with overall comprehension in Dataset 3 ($r = 0.45$). In turn, individual differences were moderately to strongly correlated with each other across all three datasets. These correlations confirm the validity of the comprehension measure as well as theoretical assumptions regarding the impact of reading skills and prior knowledge on comprehension.

Rubric scores

Appendix B provides detailed means and standard deviations for each of the rubric categories averaged across all datasets and for each dataset and text. Categories that were rarely observed among the responses (<0.1%) were excluded from Appendix B. Specifically, of 5,858 self-explanations across all datasets, only 14 were considered too short to code, 29 were entirely irrelevant, 46 included partially irrelevant statements, 6 were solely evaluative, and only 2 responses included a reference to a life event; thus, these are not included. There were also few responses that included evaluative statements (n = 56), misconceptions (n = 120), or comprehension monitoring statements (n = 91). In the context of self-explaining science texts, comprehension monitoring statements are relatively rare because the process of self-explanation inherently requires monitoring comprehension, and thus explicitly describing the process is unnatural. Approximately 24% of the responses included some text that was copied or retyped from the source text; however, only 4% (i.e., n = 251) were identified as solely including the source text and thus allotted a quality score of 0.

The rubric scores revealed that students predominantly paraphrase and bridge, with few elaborations. Because these texts are relatively challenging texts about scientific topics, students have less prior knowledge about the topics compared to, for example, narrative passages. Thus, elaborations, which rely on prior knowledge, were relatively rare, occurring in less than 9% of the responses across the datasets. This may be attributed to the fact that the texts used were relatively difficult scientific texts and thus students had too little prior knowledge to support elaborative inferences in the context of attempting to explain the meaning of the text.

The average quality score of the self-explanations was moderate (i.e., M = 1.57), which is to be expected given that the self-explanations were produced without extensive instruction and practice. Table 9 further shows that few responses received a poor quality score of 0 (i.e., 2 to 11% across the three datasets), and most were rated in quality as either fair or good.

Combinations of strategies

We examined strategy use at the response level by identifying combinations of strategies used within each response. Our objective was to characterize each self-explanation response in terms of strategy use. Thus, we examined the proportion of responses that used each strategy independently or in combination with other strategies. That is, we categorically labeled each self-explanation as (1) none, (2) paraphrase only, (3) bridge only, (4) elaboration only, (5) para-bridge (i.e., paraphrase-bridge), (6) para-elab (i.e., paraphrase-elaboration), (7) bridge-elab (i.e., bridge-elaboration), or (8) all (i.e., paraphrase-bridge-elaboration). As illustrated in Table 2, the term para-bridging should not be interpreted to reflect a particular sequence of strategy use within a given self-explanation. That is, a reader may have produced a para-bridging self-explanation by first paraphrasing a part of the sentence and then producing a bridge, or they may have generated a bridging inference first and then added a paraphrase later.

Figure 1 illustrates that participants' responses predominately comprised a combination of paraphrasing and bridging inferences across all of the datasets. Paraphrasing was the second most commonly used strategy. All other strategy combinations accounted for less than 10% of the responses in the datasets. Elaboration by itself and the combination of elaboration and bridging occurred in less than 1% of responses in any dataset. Across the datasets, participants commonly either paraphrased the target sentence, or anchored a bridging inference using a paraphrase. This finding underscores the importance of considering the role of the strategy within a response, beyond simply looking at the frequency of using a particular strategy (i.e., as in Table 9).

Relations between strategy use, comprehension test, and individual differences

Table 9 provides the means and correlations among reading behaviors and comprehension performance averaged across the three datasets, and individual differences in reading skill and

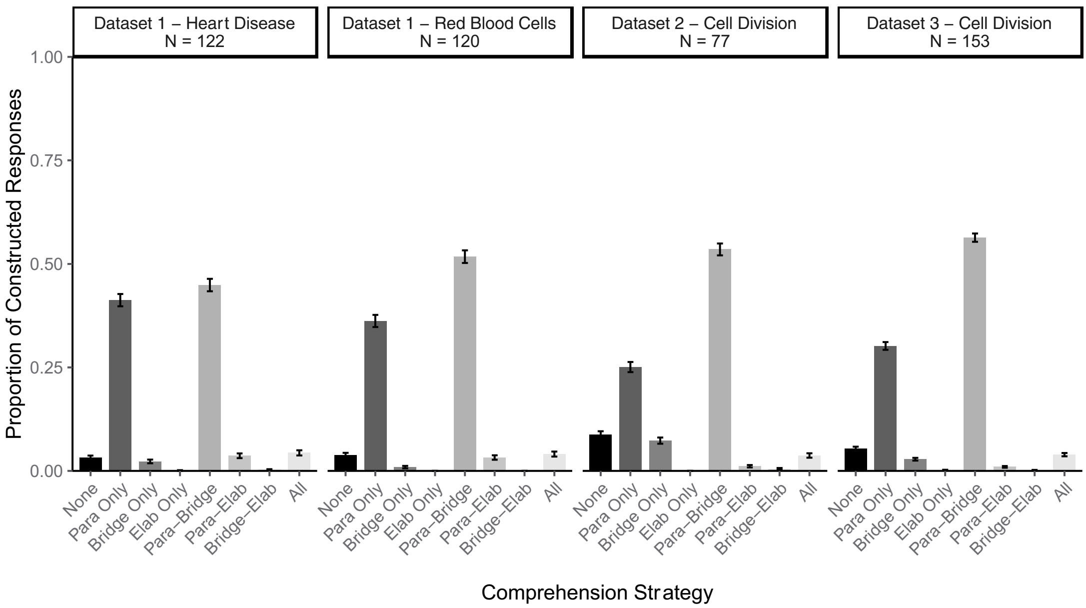

Figure 1. Frequency of combinations of comprehension strategy use per text and dataset.

prior knowledge for each dataset separately. Correlations were calculated based on participant means. Table 9 also provides the number of times that participants switched between strategies (i.e., switch) and average comprehension score. It is of note that several strategies occurred very rarely: bridging only (n = 194, M = 0.03, SD = 0.08), elaboration only (n = 6, M = 0.001, SD = 0.1), para-elab (n = 113, M = 0.02, SD = 0.07), bridge-elab (n = 13, M = 0.002, SD = 0.02), the combination of all three strategies (n = 234, M = 0.04, SD = 0.09), as well as the use of no strategy (n = 316, M = 0.05, SD = 0.13). We provide the full correlation table for clarity but emphasize that these correlations of highly skewed values should be interpreted with caution.

The correlations among the strategies are consistent with prior findings. The use of no strategy (none) and paraphrase only was negatively correlated with the use of para-bridge and the combination of all three strategies, as well as quality score, comprehension performance, and individual differences in literacy and prior knowledge. Bridges, elaborations, and bridge-elab (albeit rare) also yielded negative and low correlations with performance and individual differences. Importantly, the strategy combinations that were positively correlated with comprehension performance and individual differences in literacy and prior knowledge were inferences that were anchored by a paraphrase of the target sentence (i.e., para-bridge, para-elab, and all strategies).

As expected, the overall quality of the self-explanation was strongly correlated with overall comprehension, and more strongly associated with textbase (r = 0.49) than bridging (r = 0.38) question performance. Quality was also correlated with individual differences in reading skill (Dataset 1; r = 0.38), vocabulary knowledge (Dataset 3; r = 0.42); prior general knowledge (Dataset 1; r = 0.36; Dataset 3; r = 0.44), and prior biology knowledge (Dataset 2; r = 0.38; Dataset 3; r = 0.36).

Notably, participants' propensity to switch between strategies was negatively correlated with response quality and para-bridging, which was the most successful strategy. While switching strategies might be considered as a sign of flexibility in strategy use, here it appears to be more associated with the use of various strategies that contribute less to coherence-building, which entails making connections between the target sentence and other ideas within the text. Switching between strategies was not correlated with performance or individual differences. We further explore the nature of flexibility below.

Table 9. Correlations between strategy use and individual differences.

Variable (Proportion)	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13
1. None	0.05	0.13													
2. Paraphrase Only	0.34	0.25	−0.02												
3. Bridge Only	0.03	0.08	−0.02	−0.13**											
4. Elaboration Only	0.00	0.01	0.05	−0.03	0.18**										
5. Bridge-elab	0.00	0.02	−0.02	−0.10*	0.21**	0.51**									
6. Para-bridge	0.52	0.26	−0.40**	−0.71**	−0.18**	−0.12*	−0.07								
7. Para-elab	0.02	0.07	−0.10*	−0.19**	−0.01	−0.02	0.03	−0.17*							
8. All	0.04	0.09	−0.14**	−0.39**	0.00	0.03	0.08	−0.02	0.34**						
9. Switch Score	0.48	0.21	0.11*	0.08	0.20**	0.11*	0.12**	−0.37**	0.30**	0.24**					
10. Quality	1.62	0.57	−0.54**	−0.54**	−0.12*	−0.07	−0.01	0.64**	0.18**	0.39**	−0.13**				
11. Comprehension Performance	0.47	0.23	−0.34**	−0.31**	−0.14**	−0.01	−0.05	0.37**	0.13**	0.30**	−0.05	0.50**			
12. Textbase Comprehension Performance	0.53	0.28	−0.33**	−0.28**	−0.15**	−0.03	−0.08	0.36**	0.15**	0.24**	−0.05	0.49**	0.91**		
13. Bridging Comprehension Performance	0.39	0.25	−0.26**	−0.29**	−0.10*	0.03	0.00	0.31**	0.06	0.29**	−0.03	0.38**	0.86**	0.60**	
14. Reading Skill Dataset 1	0.57	0.22	−0.24**	−0.21**	−0.21**	−0.10	−0.10	0.23**	0.09	0.24**	0.01	0.38**	0.53**	0.45**	0.49**
15. Vocabulary Dataset 3	0.68	0.19	−0.32**	−0.29**	−0.11	0.13	0.04	0.35**	0.16	0.36**	−0.16*	0.42**	0.34**	0.37**	0.27**
16. General Knowledge Dataset 1	0.57	0.15	−0.28**	−0.20**	−0.16*	−0.09	−0.06	0.18**	0.13	0.30**	0.10	0.36**	0.53**	0.46**	0.46**
17. General Knowledge Dataset 3	0.58	0.18	−0.22**	−0.38**	−0.05	0.13	0.01	0.37**	0.13	0.36**	−0.17*	0.44**	0.39**	0.41**	0.32**
18. Biology Knowledge Dataset 2	0.38	0.16	−0.38**	−0.35**	0.18	N/A	0.18	0.35**	0.04	0.24*	−0.09	0.38**	0.64**	0.63**	0.58**
19. Biology Knowledge Dataset 3	0.42	0.15	−0.20*	−0.30**	−0.03	0.23**	0.16	0.28**	0.18*	0.32**	−0.15	0.36**	0.45**	0.45**	0.37**

"Bridge-elab" refers to the combination of bridging and elaboration. "Para-bridge" refers to the combination of paraphrasing and bridging. "Para-elab" refers to the combination of paraphrasing and elaboration. "All" refers to the combination of paraphrasing, bridging, and elaboration. The ranges are 0–1 for all variables except for quality, which ranges from 0 to 3.

*p < .05. **p < .01.

Examining transitions with Markov chains

To better understand the nature of students' transitions between strategies, a Markov chain visualization was calculated. In Figure 2, each node represents the overall frequency of a strategy (no strategies, paraphrasing in isolation, bridging in isolation, the combination of paraphrasing and elaborating, the combination of paraphrasing and bridging, and the combination of paraphrasing, bridging, and elaborating). The base rate of the strategy, meaning how often that strategy combination was used overall, is indicated as a percentage within each node. Consistent with the aggregate analysis above, these visualizations highlight the finding that participants primarily used para-bridging, as well as solely paraphrasing. Elaboration (n = 6) and bridge-elab (n = 13) were excluded from the Markov analysis, as there were very few instances of them. Participants' probability of continuing to use the same strategy or switching to a different strategy are displayed by the widths of the arrows; probabilities above 20% are darker and labeled. The transitional probabilities can be contrasted to the base rate of each strategy to describe whether participants are switching randomly or forming patterns.

Overall, when using a strategy in one self-explanation, participants tended to continue to use the same strategy in the subsequent self-explanation rather than switching to a different strategy. For example, if a participant combined paraphrasing and bridging in one self-explanation, there was a 64% chance that their next self-explanation would be a para-bridge. In comparison, overall para-bridging only occurred 53% of the time. Thus, whether a participant para-bridged in a self-explanation is related to whether they para-bridged previously. When participants paraphrased in isolation, they continued to do so about half the time, although overall paraphrasing only occurred 32% of the time. Thus, participants were more likely to paraphrase if they had paraphrased the previous sentence. Very few switches between strategies occurred that were not to or from paraphrasing or para-bridging. When bridging in isolation, para-elaborating, or using all three strategies, participants were most likely to switch to para-bridging (40%-58% probability). When participants used no strategies at all, they were most likely to switch to paraphrasing in isolation (38%). These transitions suggest that most participants were consistently using paraphrasing or para-bridging, suggesting that their self-explanations were

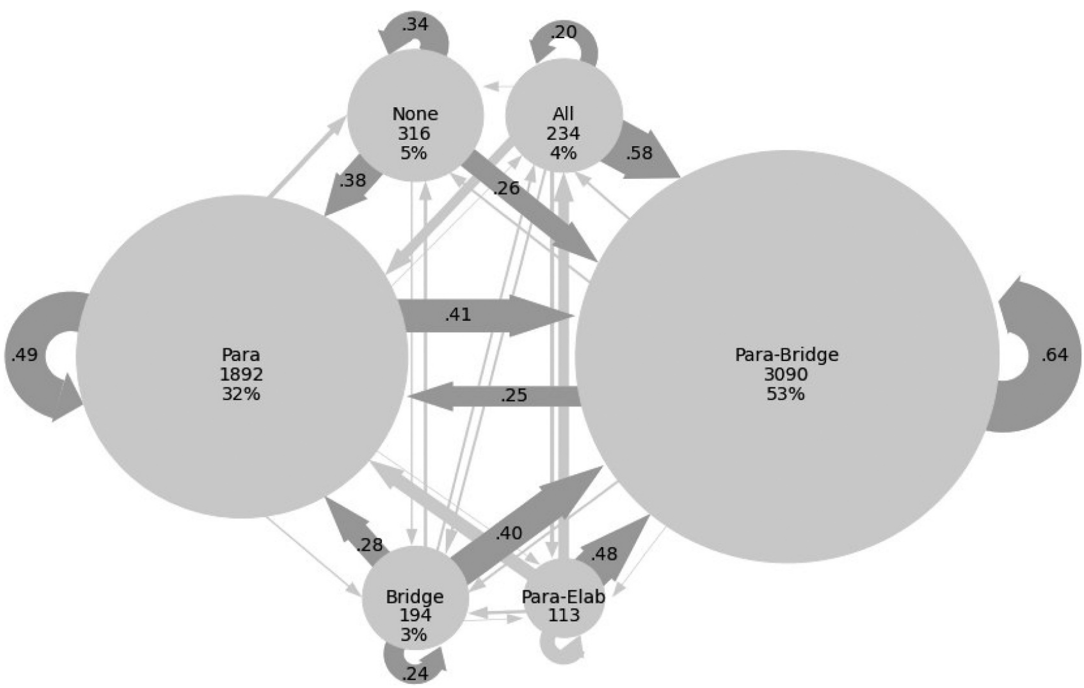

Figure 2. Markov chain diagram demonstrating probabilities of switching between different combinations of strategies.

anchored in the text. If a participant chose to use a different strategy combination for a given self-explanation, they likely returned to paraphrasing or para-bridging in the next self-explanation.

Predicting comprehension

We conducted a linear mixed model analyses using strategy combinations to predict text comprehension. All analyses were performed using R Statistical Software (v4.2.2; R Core Team, 2023). The multilevel regression analyses were conducted using maximum likelihood estimation through the lme4 package, version 1.1-30 (Bates et al., 2015) and the sjPlot package, version 2.8.12 (Lüdecke, 2022). The intraclass correlation was .117, which suggests that, before adding predictors, about 12% of variance in text comprehension can be accounted for by clustering in the data. Specifically, 3% of the variance in text comprehension was due to text clustering, and about 9% of the variance was due to dataset clustering, with the remaining 88% due to individual factors. As such, we accounted for clustering by text and by dataset in our analyses. Our analysis was restricted to the strategies (and strategy combinations) that demonstrated moderate correlations with comprehension (see Table 10). We used a random intercepts model with no strategies, paraphrasing, para-bridging, and all strategies as individual-level predictors, and text and dataset as clustering variables. The analysis included 471 participants, clustered within three texts and three datasets.

As would be expected, using no strategies negatively predicted comprehension, such that a participant who consistently used no strategies would be estimated to score 25 percentage points worse on the comprehension measure, compared to a student who always used at least one strategy ($t = -2.332$, $df = 463.68$, $p = .020$). Paraphrasing alone did not significantly account for variance in comprehension. Para-bridging ($t = 3.57$, $df = 464.70$, $p < .001$) and using all three strategies ($t = 4.75$, $df = 464.00$, $p < .001$) both contributed positively to comprehension.[4] A student who continually para-bridged would be expected to score 31 percentage points above students who never para-bridged. A student who continually used all three strategies would be estimated to score 70 percentage points above students who never combined all three strategies. Thus, anchoring a bridging inference with a paraphrase (with or without an elaborative inference) was the key strategy contributing to enhanced comprehension. Collectively, these fixed effects explained about 24% of the variance in comprehension. The remaining unexplained variance can be divided into clustering by text (3%), clustering by dataset (8%), and residual variance due to individual factors (64%). The reduction in residual variance from 88% in the model without predictors to 64% in the final model suggests that strategy use primarily explained individual-level variance, not text- or dataset-level variance.[5]

Discussion

Most studies using constructed responses (e.g., self-explanations, think-alouds) have either emphasized the identification of individual strategies (e.g., McMaster et al., 2012; McNamara, 2004) or

Table 10. Regression analysis: predicting text comprehension from strategy use.

Variable	*Estimate*	*SE*	*df*	*t*	*p*
Fixed Effects					
Intercept	0.28	0.10	29.70	2.86	.008**
No Strategies	−0.25	0.11	463.68	−2.33	.02*
Paraphrasing Only	−0.00	0.09	464.73	−0.00	1.00
Para-bridging	0.31	0.09	464.70	3.57	<.001***
All strategies	0.70	0.15	464.00	4.75	<.001***
Variance Components					
Text Variance	.002	.045			
Dataset Variance	.005	.071			
Residual Variance	.039	.197			

$R^2 = 0.24$; Model: TextComp ~ None + Para + ParaBridge + All + (1|Text) + (1|Dataset); "Para-bridging" refers to the combination of paraphrasing and bridging; *$p < .05$. **$p < .01$. ***$p < .001$.

a global assessment of a general strategy (Coté & Goldman, 1999). As such, there is a dearth of research on how multiple strategies are coordinated to support comprehension in the context of constructed responses. While some researchers have made claims about how strategies are combined to support meaning making and more specifically that paraphrasing is important for bridging and elaboration (e.g., McNamara, 2004), more research is needed to test these claims. To this end, this study used a large data set to explore the comprehension strategies used by high school and college students as they read science texts. More specifically, we expanded on existing research in constructed response coding to examine how readers combine and coordinate their strategy use and the extent to which such behaviors are related to individual differences and comprehension outcomes.

We found relative stability across datasets, with results revealing that analyzing combinations of strategies afforded a nuanced understanding of how readers engage in coherence-building. In particular, more skilled and successful readers tended to engage in *anchored* inferencing that involved bridging and elaborations that were combined with paraphrases. By contrast, we found little evidence that strategy flexibility was related to reader differences or comprehension.

Exploring combinations of strategies

Our correlational analyses were largely in line with prior literature that shows a positive relationship between self-explanation quality and comprehension (e.g., Chi et al., 1989, 1994; Magliano et al., 1999a). There were moderate to strong, positive correlations between overall self-explanation quality and comprehension outcomes (*r*s from 0.38 to 0.50). There were also moderate to strong correlations between comprehension outcomes and readers' individual differences in literacy and knowledge (*r*s from 0.45 to 0.64). A novel contribution of our study was to examine how readers may use a single strategy or combinations of strategies when generating a self-explanation. Our coding scheme allowed us to categorize self-explanations from having no evidence of the focal comprehension strategies to having evidence of all three strategies.

When examining the strategies in isolation, it appeared that there was a substantial amount of paraphrasing in the self-explanations, which is usually interpreted as evidence of less coherence-building. However, our analyses indicated that much of the paraphrasing in the self-explanations was done in combination with more effective inferential strategies (e.g., bridging, elaboration). The frequent use of *para-bridging* is likely a useful strategy for self-explanation because combining paraphrasing and bridging works to establish how the sentence that was just read is related to the prior discourse context. That is, readers were anchoring their inferences to the text. Indeed, the correlational analyses showed that paraphrasing without additional strategies was negatively related to comprehension, but that para-bridging, para-elaborating, and using all three strategies were positively related to comprehension. While it is possible to treat paraphrasing and bridging as distinct categories in a coding system, the results of the present study suggest that effective bridging may require paraphrasing. As such, future researchers may consider para-bridging as a distinct category in their coding system.

Exploring patterns and flexibility of strategy use

A second contribution of this study was to explore comprehension strategy use beyond proportions and aggregates. We investigated patterns and flexibility in two ways: switching scores and Markov Chains. Given the limited work on flexibility, there were competing hypotheses regarding whether switching strategies would be indicative of stronger or weaker coherence-building. Various analytic approaches with which we have explored with these types of datasets (e.g., cluster analyses; Butterfuss et al., 2022; Christhilf et al., 2022) have led to similar conclusions: the primary predictor of successful text comprehension is the combined use of paraphrasing and bridging. We do not observe a wide variety of strategy use: most self-explanations are either solely paraphrases or para-bridges.

Further, our findings indicated that the switching score was not related to comprehension and was, in fact, negatively related to the use of more effective strategies (e.g., para-bridging). In the context of the targeted studies, these results suggest that strategy switching may be relatively unsystematic and that switching may reflect increased use of less effective comprehension strategies, such as solely paraphrasing. For our data, forming discrete reader profiles based on flexibility of strategy use would be less accurate than considering the frequency with which readers para-bridge as a continuous variable. This is in line with past research that suggests the importance of connecting ideas across the text is essential to successful comprehension (Magliano et al., 2011; Singer et al., 1992).

Similarly, the Markov chains showed that participants tended to repeatedly use the same strategy. In almost all cases the predominant transition was back to para-bridging. The only exception to this was the transition after a self-explanation that had no strategies. In this case, participants were most likely to switch to paraphrasing in the subsequent self-explanation. This pattern of switching from no strategies to paraphrases may be indicative of struggling readers, either due to a need for additional training or due to particularly difficult sections of text. In these circumstances, readers may not know how to enact any comprehension strategies initially, but then may switch to simply paraphrasing to ground a basic understanding of the text content. Our analyses provide some support for this, as para-bridging and the combination of all strategies were positively correlated with comprehension performance, whereas the propensity to not engage in any strategies was negatively correlated with comprehension outcomes. Thus, examinations of strategy combinations, such as the ones present in the current study, provide further evidence that constructed responses such as self-explanation have strong utility in revealing *how* comprehension is achieved by readers.

While we did not explore profiles in the present study, there would likely be two that emerged, specifically paraphrasers and para-bridgers, the former of which has been reported in other studies (Carlson et al., 2014; Karlsson et al., 2018; McMaster et al., 2012; Rapp et al., 2007). Nonetheless, how do we reconcile the results of this study with perspectives that flexibility in strategy use is important for comprehension (Cromley & Wills, 2014; McNamara & Magliano, 2009a)? First, flexibility in strategy use could be more related to task-level strategies that arise from metacognitive states that are indicative of comprehension challenges, such as rereading, question asking, or summarizing (e.g., McNamara, 2007; Pressley & Afflerbach, 1995). Self-explanation is in itself a task-level strategy, and thus, flexibility may be reflected by the choice to self-explain or not, ask questions or not, and so on, rather than switching between processing strategies in the context of self-explanation. Skilled readers may more flexibly adopt task-level strategies to meet their goals in the context of the challenges presented by the text. One possible way to further explore flexibility in processing strategies is to vary task-level strategies that affect how processing strategies are used to accomplish those tasks (Van den Broek et al., 2001).

Second, self-explanation may inherently benefit from anchored bridges, particularly in the context of science texts wherein readers likely lack sufficient knowledge to generate elaborative inferences. The success and predominant use of para-bridging as a processing strategy to build coherence may be driven by the task (self-explanation) and the nature of the texts targeted in this study: The self-explanations were generated solely in the context of science texts. The extent to which the target sentences required coherence-building (e.g., the target sentences had causal relationships with prior discourse sentences) likely impacted the results reported in this study (Trabasso et al., 1989). The comprehension processes that are recruited to build and maintain coherence change from moment to moment as the properties of the text change across sentences (e.g., Van den Broek et al., 2015, 1999). Indeed, McNamara and Magliano (2009a) argued that flexibility within the context of self-explanation is in part related to the semantic relationships across sentences that change from sentence to sentence (e.g., argument overlap, new arguments, causal relationships). They posited that the use strategies that comprise self-explanation should vary as a function of text features, such as causal cohesion and the introduction of new topics. The texts use in these studies combined with the methodological decision to have participants produce self-explanations at a limited number of sentences (e.g., 9 sentences) does not afford exploring flexibility of processing strategies as a function of text features.

Comprehension processes also vary as a function of readers' goals (e.g., Van den Broek et al., 2001, 1999) as well as disciplines (Goldman & Snow, 2015) and text genres (Clinton et al., 2020; McNamara et al., 2012). Comprehension of science texts is quite different from other genres, such as science and history; thus, we would expect strategies to vary across these types of texts. Narratives call for understanding characters, plots, and events, which are often presented out of temporal order. History texts similarly call for world knowledge, as well as developing an understanding of complex roles, protagonists and antagonists, combined with multiple viewpoints (sources) on what might be an accurate portrayal of history. As such, elaborations are more prone to come into play, and potentially, readers may benefit more from switching between strategies. Likewise, narratives and history texts may vary more in the demands of the text. For example, some narratives transition between information delivery (history, science), dialogs, and storytelling.

Skilled readers in turn are likely to respond to variations in text in tune with their skills, experience, and knowledge. Investigating these differences is critical to better understanding how coherence is constructed across multiple contexts and textual demands. Investigating processing strategies as a function of variations in linguistic and semantic relationships in text calls for additional studies that include a larger set of texts that vary more widely (e.g., narratives, history, science) as well as multiple texts per participant, which would afford examining how and whether readers adapt strategies depending on the nature of the text as well as the features of the sentence (see e.g., Magliano et al., 1999b; Ray & Magliano, 2015; Trabasso & Suh, 1993).

Limitations

Of course, there are multiple limitations of this study. Among them, first, we only examined high school and college students, which inherently limits the generalizability of our claims. Similar questions should be examined among younger and older students in future research. Second, this study focused solely on constructed responses in the context of self-explanation. While this focus was motivated by the inherent affordances of self-explanation, exploring the extent that these results are replicated in the context of other tasks merits consideration. Third, our coding rubric defined bridging by evidence of semantic content from outside the current sentence context but not the degree to which a reader makes explicit the connection between the current and previous content. That is, our rubric did not distinguish between responses in which participants were simply restating the text (i.e., reproduces content, but does not show evidence of integrating across the content) or demonstrated integration across the content. As such, some responses that were coded as constituting the highest level of bridging may reproduce content without explicitly describing the relation between the ideas. In future work, it would be of value to evaluate the presence or degree of explicit connections that readers make and how those connections relate to comprehension outcomes.

Notably, the relations reported here are also constrained by the comprehension measure itself. Answering textbase questions requires building a coherent mental model that can be accessed after reading the questions and paraphrasing and bridging likely reflect the coherence of the mental model. The bridging inference questions target the extent that readers represent relationships across discourse content, which similarly requires readers to build a coherent mental model. This relative *task appropriateness* does not, however, explain the patterns: Students did not know a priori that the questions would tap into their bridging inferences as opposed to elaborations. They were only informed that their task was to construct a coherent understanding of the text.

Conclusions

Comprehension strategies are usually examined at the aggregate, which ignores the notion that readers are likely to coordinate and combine strategies across the text (McNamara & Magliano, 2009a). Our results demonstrate that studies that use constructed responses but rely on coding schemes or analyses that focus strategies in isolation are not adequately conveying how strategies support coherence-building. As such, we encourage researchers to take this under consideration when they design future studies. Doing so will have important implications for theory testing and practical applications, such as comprehension strategy training.

Self-explanation reveals strategies inherent to coherence-building (e.g., McNamara & Magliano, 2009a), and thus informs our understanding of how students leverage comprehension strategies while reading challenging texts. This is important both to inform theoretical accounts of comprehension (Graesser et al., 1994; Kintsch, 1988; Myers & O'Brien, 1998), as well as individualized interventions to help students improve their comprehension skills by targeting specific comprehension issues (McNamara, 2004; McNamara et al., 2004). We argued that there is a dearth in the literature regarding how strategies are coordinated in the context of challenging text. The present study provided important insights regarding how strategies support coherence-building processes and comprehension. Successful comprehenders do not produce inference strategies in isolation.

While theories of comprehension have generally assumed that strategies are coordinated, few studies have adopted analytic approaches to explore this claim. This study shows the importance of anchoring inference strategies in the immediate discourse context, and paraphrasing and bridging was the predominant way that this was achieved. Although we used multiple data sets, we encourage other researchers to explore the extent that these findings replicate across populations, age groups, text genres, and instructions. Doing so will both inform theory and practice regarding the nature of comprehension and how to help readers of all ages accomplish it, and in particular when they are faced with text that they find challenging. In conclusion, we encourage researchers and practitioners to carefully consider the importance of encouraging readers to engage in meaning making strategies that ground the ideas in words that they understand (i.e., paraphrasing) and anchoring their bridges that connect ideas from across the text.

Notes

1. In recent work, we have adopted the term *constructed response* as a general term to refer to any concurrent protocol instructions, including prompting to think aloud or instructions to self-explain, attend to sources, and so on. We use the term constructed response instead of verbal protocol because verbal protocol 1) is often used to refer to "unbiased" thinking aloud (e.g., Ericsson & Simon, 1993; Pressley & Afflerbach, 1995) and 2) can imply spoken responses, whereas constructed responses can have a variety of instructions and registers.
2. https://osf.io/tm5vy/?view_only=38167d5ca5614c14aeed63bb7847b973
3. Participants from Dataset 2 (Allen et al., 2017) self-explained a practice text on the topic of heart disease prior to self-explaining Cell Division. However, they did not receive feedback on the quality of their self-explanations. We chose not to include the Heart Disease self-explanations in this study because it was presented as a practice text (not counterbalanced), and there were no comprehension questions associated with the text.
4. We also conducted this analysis restricted to the strategies that were normally distributed (i.e., paraphrasing and para-bridging). In this restricted analysis, paraphrasing alone was negatively related to text comprehension ($t = -1.47$, $df = 466.42$, $p = .007$), while para-bridging continued to significantly predict comprehension ($t = 5.01$; $df = 465.99$, $p < .001$), with the model accounting for abut 15% of the variance in comprehension.
5. We had initially conducted a cluster analysis to examine the relation of strategy profiles to comprehension (Butterfuss et al., 2022). However, the clusters did not provide additional insights beyond what was apparent from the linear analyses we are presenting here. Essentially, the clusters simply reveal the tendency to para-bridge or paraphrase, which is a linear relationship.

   This finding is likely attributable to most self-explanations being categorized as either paraphrases or para-bridges.

Acknowledgments

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A190063 to Arizona State University. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education. The collection of the data associated with the three datasets were funded by R305A130124 and R305A120707, R305A190063. This study was not preregistered. The data, appendices, and analytic code are available at McNamara, D., Newton, N., Christhilf, K., McCarthy, K.S., Magliano, J.P., & Allen, L. (2023). Anchoring Your Bridge: The Importance of Paraphrasing to Inference Making in Self-Explanations [Data set, appendices, and code]. OSF. https://osf.io/tm5vy/

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Institute of Education Sciences [R305A130124, R305A120707, R305A190063].

References

Alexander, P. A., Kulikowich, J. M., & Jetton, T. L. (1994). The role of subject-matter knowledge and interest in the processing of linear and nonlinear texts. *Review of Educational Research*, *64*(2), 201–252. https://doi.org/10.2307/1170694

Allen, L. K., Perret, C., Guerrero, T., & McNamara, D. S. (2017, July). *You talkin' to me? The role of audience in the generation of explanations* [Poster presentation]. Annual Meeting of the Society for Text and Discourse, Philadelphia, PA.

Anmarkrud, O., Braten, I., & Stromso, H. I. (2014). Multiple-documents literacy: Strategic processing, source awareness, and argumentation when reading multiple documents. *Learning and Individual Differences*, *30*, 64–76. https://doi.org/10.1016/j.lindif.2013.01.007

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

Best, R., Ozuru, Y., & McNamara, D. S. (2004, October). Self-explaining science texts: Strategies, knowledge, and reading skill. In *Embracing Diversity in the Learning Sciences: Proceedings of the Sixth International Conference of the Learning Sciences* (p. 89). Psychology Press.

Butterfuss, R., Christhilf, K., Newton, N. N., McCarthy, K. S., Magliano, J. P., Allen, L. K., & McNamara, D. S. (2022, July). *Understanding flexible use of comprehension strategies: A cluster approach* [Poster presentation]. Annual Meeting of the Society for Text and Discourse (ST&D), Virtual Conference.

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

Chi, M. T. H., Bassok, M., Lewis, M., Reimann, P., & Glasser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, *13*(2), 145–182. https://doi.org/10.1207/s15516709cog1302_1

Chi, M. T. H., De Leeuw, N., Chiu, M.-H., & Lavancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, *18*(3), 439–477.

Chi, M. T. H., & VanLehn, K. A. (1991). The content of physics self-explanations. *The Journal of the Learning Sciences*, *1*(1), 69–105. https://doi.org/10.1207/s15327809jls0101_4

Christhilf, K., Newton, N. N., Butterfuss, R., McCarthy, K. S., Allen, L. K., Magliano, J. P., & McNamara, D. S. (2022). Using Markov models and random walks to examine strategy use of more or less successful comprehenders. In *Proceedings of the 15th International Conference on Educational Data Mining*. Durham, UK.

Clinton, V., Taylor, T., Bajpayee, S., Davison, M. L., Carlson, S. E., & Seipel, B. (2020). Inferential comprehension differences between narrative and expository texts: A systematic review and meta-analysis. *Reading and Writing*, *33*(9), 2223–2248. https://doi.org/10.1007/s11145-020-10044-2

Corbett, A. T., & Anderson, J. R. (1995). Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction*, *4*, 253–278. https://doi.org/10.1007/BF01099821

Coté, N., & Goldman, S. R. (1999). Building representations of informational text: Evidence from children's think-aloud protocols. In H. Van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 169–193). Erlbaum.

Coté, N., Goldman, S. R., & Saul, E. U. (1998). Students making sense of informational text: Relations between processing and representation. *Discourse Processes*, *25*(1), 1–53. https://doi.org/10.1080/01638539809545019

Creer, S. D., McCarthy, K. S., Perret, C. A., McNamara, D. S., Magliano, J. P., & Allen, L. K. (2020, November). *Cohesion and coherence: Using Natural Language Processing techniques to examine successful comprehension* *processes* [Paper presentation]. 50th Annual Conference of The Society for Computation in Psychology (SCiP), Virtual Conference.

Cromley, J. G., & Wills, T. W. (2014). Flexible strategy use by students who learn much versus little from text: Transitions within think-aloud protocols. *Journal of Research in Reading*, *39*(1), 50–71. https://doi.org/10.1111/1467-9817.12026

Cumming, G. (2014). The New Statistics: Why and How. *Psychological Science*, *25*(1), 7–29. https://doi.org/10.1177/0956797613504966

D'Mello, S., & Graesser, A. (2010). Modeling cognitive-affective dynamics with Hidden Markov Models. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *32*(32), 2721–2726.

Dahl, A. C., Carlson, S. E., Renken, M., McCarthy, K. S., & Reynolds, E. (2021). Materials matter: An exploration of text complexity and its effects on middle school readers' comprehension processing. *Language, Speech, and Hearing Services in Schools*, *52*(2), 702–716. https://doi.org/10.1044/2021_LSHSS-20-00117

Dai, D. Y., & Wang, X. (2007). The role of need for cognition and reader beliefs in text comprehension and interest development. *Contemporary Educational Psychology*, *32*(3), 332–347. https://doi.org/10.1016/j.cedpsych.2006.05.002

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data. A bradford book*. The MIT Press.

Farley, M. J., & Elmore, P. B. (1992). The relationship of reading comprehension to critical thinking skills, cognitive ability, and vocabulary for a sample of underachieving college freshmen. *Educational and Psychological Measurement*, *52*(4), 921–931. https://doi.org/10.1177/00131644920520040

Geyer, C. J. (1992). Practical Markov chain Monte Carlo. *Statistical Science*, *7*(4), 473–483.

Goldman, S. R., Braasch, J. L. G., Wiley, J., Graesser, A. C., & Brodowinska, K. (2012). Comprehending and learning from internet sources: Processing patterns of better and poorer learners. *Reading Research Quarterly*, *47*(4), 356–381.

Goldman, S. R., & Saul, E. U. (1990). Flexibility in text processing: A strategy competition model. *Learning and Individual Differences*, *2*(2), 181–219. https://doi.org/10.1016/1041-6080(90)90022-9

Goldman, S. R., & Snow, C. E. (2015). Adolescent literacy: Development and instruction. In A. Pollatsek & R. Treiman (Eds.), *The Oxford handbook of reading* (pp. 463–478). Oxford University Press.

Graesser, A. C., & Clark, L. F. (1985). The generation of knowledge-based inferences during narrative comprehension. *Advances in Psychology*, *29*, 53–94.

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

Haynes, J. E., & Fillmer, H. T. (1984). Paraphrasing and reading comprehension. *Reading World*, *24*(1), 76–79.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*(1), 122–149. https://doi.org/10.1037/0033-295X.99.1.122

Karlsson, J., van den Broek, P., Helder, A., Hickendorff, M., Koornneef, A., & van Leijenhorst, L. (2018). Profiles of young readers: Evidence from thinking aloud while reading narrative and expository texts. *Learning and Individual Differences*, *67*, 105–116. https://doi.org/10.1016/j.lindif.2018.08.001

Kintsch, W. (1988). The use of knowledge in discourse processing: A construction-integration model. *Psychological Review*, *95*(2), 163–182. https://doi.org/10.1037/0033-295X.95.2.163

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

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*(5), 363–394. https://doi.org/10.1037/0033-295X.85.5.363

Koslowski, B., & Masnick, A. (2002). The development of causal reasoning. In U. Goswami (Ed.), *Handbook of childhood cognitive development* (pp. 257–281). Blackwell.

Kurby, C. A., Magliano, J. P., Dandotkar, S., Woehrle, J., Gilliam, S., & McNamara, D. S. (2012). Changing how students process and comprehend texts with computer-based self-explanation training. *Journal of Educational Computing Research*, *47*(4), 429–459. https://doi.org/10.2190/EC.47.4.e

Lombrozo, T. (2006). The structure and function of explanations. *Trends in Cognitive Sciences*, *10*(10), 464–470. https://doi.org/10.1016/j.tics.2006.08.004

Long, D. L., Seely, M. R., & Oppy, B. J. (1999). The strategic nature of less skilled readers' suppression problems. *Discourse Processes*, *27*(3), 281–302. https://doi.org/10.1080/01638539909545064

Lüdecke, D. (2022). *sjPlot: Data visualization for statistics in social science*. R package version 2.8.12. https://CRAN.R-project.org/package=sjPlot

MacGinitie, W. H., & MacGinitie, R. K. (1989). Gates-MacGinitie reading tests, third edition. In *The reading teacher* (Vol. 43, No. 3, pp. 256–258). International Literacy Association and Wiley. https://www.jstor.org/stable/20200351

Magliano, J. P., & Graesser, A. C. (1991). A three-pronged method for studying inference generation in literary text. *Poetics*, *20*(3), 193–232. https://doi.org/10.1016/0304-422X(91)90007-C

Magliano, J. P., Higgs, K., Santuzzi, A., Tonks, S., O'Reilly, T., Sabatini, J., Feller, D., Kopatich, R., Ray, M., & Parker, C. (2020). Testing the inference mediation hypothesis in a post-secondary context. *Contemporary Educational Psychology*, *61*, 101867. https://doi.org/10.1016/j.cedpsych.2020.101867

Magliano, J. P., & Millis, K. K. (2003). Assessing reading skill with a think-aloud procedure and latent semantic analysis. *Cognition and Instruction*, *21*(3), 251–283. https://doi.org/10.1207/S1532690XCI2103_02

Magliano, J. P., Millis, K. K., Levinstein, I., & Boonthum, C. (2011). Assessing comprehension during reading with the Reading Strategy Assessment Tool (RSAT). *Metacognition and Learning*, *6*(2), 131–154. https://doi.org/10.1007/s11409-010-9064-2.

Magliano, J. P., Trabasso, T., & Graesser, A. C. (1999a). Strategic processes during comprehension. *Journal of Educational Psychology*, *91*(4), 615–629. https://doi.org/10.1037/0022-0663.91.4.615

Magliano, J. P., Zwaan, R. A., & Graesser, A. C. (1999b). The role of situational continuity in narrative understanding. In S. R. Goldman & H. van Oostendorp (Eds.), *The construction of mental representation during reading* (pp. 219–245). Erlbaum.

McCarthy, K. S., Kopp, K. J., Allen, L. K., & McNamara, D. S. (2018). Methods of studying text: Memory, comprehension, and learning. In H. Otani & B. Schwarz (Eds.), *Research methods in human memory* (pp. 104–124). Routledge.

McCarthy, K. S., Magliano, J. P., Snyder, J., Kenney, E., Perret, C. A., Newton, N. N., Knezevic, M., Allen, L. K., & McNamara, D. S (2021). Quantified qualitative analysis: Rubric development and inter-rater reliability as iterative design. In *Proceedings of the Annual Meeting of the International Society of the Learning Sciences (ISLS)*. Bochum, Germany (Virtual).

McCarthy, K. S., & McNamara, D. S. (2021). The multidimensional knowledge in text comprehension framework. *Educational Psychologist*, *56*(3), 196–214. https://doi.org/10.1080/00461520.2021.1872379

McCrudden, M., Huynh, L., Lyu, B., & Kulikowich, J. (2021). Bridging inferences and learning from multiple complementary texts. *Discourse Processes*, *58*(5–6), 529–548. https://doi.org/10.1080/0163853X.2021.1924586

McMaster, K. L., van den Broek, P., Espin, C. A., White, M. J., Rapp, D. N., Kendeou, P., Bohn-Gettler, C., & Carlson, S. (2012). Making the right connections: Differential effects of reading intervention for subgroups of comprehenders. *Learning and Individual Differences*, *22*(1), 100–111. https://doi.org/10.1016/j.lindif.2011.11.017

McNamara, D. S. (2001). Reading both high-coherence and low-coherence texts: Effects of text sequence and prior knowledge. *Canadian Journal of Experimental Psychology*, *55*(1), 51–62. https://doi.org/10.1037/h0087352

McNamara, D. S. (2004). SERT: Self-explanation reading training. *Discourse Processes*, *38*(1), 1–30. https://doi.org/10.1207/s15326950dp3801_1

McNamara, D. S. (Ed.). (2007). *Reading comprehension strategies: Theory, interventions, and technologies*. Erlbaum.

McNamara, D. S. (2009). The importance of teaching reading strategies. *Perspectives on Language and Literacy*, *35*, 34–40.

McNamara, D. S., Boonthum, C., Kurby, C. A., Magliano, J., Pillarisetti, S., & Bellissens, C. (2009). Interactive paraphrasing training: The development and testing of an iSTART module. In V. Dimitrova, R. Mizoguchi, B. du Boulay, & A. C. Graesser (Eds.), *Artificial intelligence in education; building learning systems that care; from knowledge representation to affective modeling* (pp. 181–188). IOS Press.

McNamara, D. S., Graesser, A. C., & Louwerse, M. M. (2012). Sources of text difficulty: Across genres and grades. In J. P. Sabatini, E. Albro, & T. O'Reilly (Eds.), *Measuring up: Advances in how we assess reading ability* (pp. 89–116). R&L Education.

McNamara, D. S., Graesser, A. C., McCarthy, P., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge University Press.

McNamara, D. S., Levinstein, I. B., & Boonthum, C. (2004). iSTART: Interactive strategy trainer for active reading and thinking. *Behavioral Research Methods, Instruments, & Computers*, *36*(2), 222–233. https://doi.org/10.3758/BF03195567

McNamara, D. S., & Magliano, J. P. (2009a). Self-explanation and metacognition: The dynamics of reading. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 60–81). Erlbaum.

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

Millis, K. K., Long, D. S., Magliano, J. P., & Wiemer, K. (Eds.). (2018). *Deep comprehension: Multidisciplinary approaches to understand, enhancing, and promoting deep comprehension*. Routledge.

Millis, K., Magliano, J., & Todaro, S. (2006). Measuring discourse-level processes with verbal protocols and latent semantic analysis. *Scientific Studies of Reading*, *10*(3), 225–240. https://doi.org/10.1207/s1532799xssr1003_2

Millis, K., Magliano, J., Wiemer-Hastings, K., Todaro, S., & McNamara, D. S. (2007). Assessing and improving comprehension with latent semantic analysis. In T. K. Landauer, D. S. McNamara, S. Dennis, & W. Kintsch (Eds.), *Handbook of latent semantic analysis* (pp. 207–225). Erlbaum.

Myers, J. L., & O'Brien, E. J. (1998). Accessing the discourse representation during reading. *Discourse Processes*, *26*(2–3), 131–157. https://doi.org/10.1080/01638539809545042

Nagy, W. (2007). Metalinguistic awareness and the vocabulary-comprehension connection. In R. K. Wagner, A. E. Muse, & K. R. Tannenbaum (Eds.), *Vocabulary acquisition: Implications for reading comprehension* (pp. 52–77). Guilford Press.

Nguyen, H., Hou, X., Stamper, J., & McLaren, B. (2020). Moving beyond test scores: Analyzing the effectiveness of a digital learning game through learning analytics. In A. N. Rafferty, J. Whitehill, C. Romero, & V. Cavalli-Sforza (Eds.), *Proceedings of The 13th International Conference on Educational Data Mining (EDM '20)* (pp. 487–495). Virtual Conference. https://educationaldatamining.org/edm2020

O'Brien, E. J., Shank, D. M., Myers, J. L., & Rayner, K. (1988). Elaborative inferences during reading: Do they occur on-line?. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(3), 410–420. https://doi.org/10.1037/0278-7393.14.3.410

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, *1*(2), 117–175. https://doi.org/10.1207/s1532690xci0102_1

Park, J., Park, C., Jung, H., & Kim, D. (2020). Promoting case indexing in case library learning: Effects of indexing prompts on self-explanation and problem solving. *Journal of Computer Assisted Learning*, *36*(5), 656–671. https://doi.org/10.1111/jcal.12435

Potts, G. R., Keenan, J. M., & Golding, J. M. (1988). Assessing the occurrence of elaborative inferences: Lexical decision versus naming. *Journal of Memory and Language*, *27*(4), 399–415. https://doi.org/10.1016/0749-596X(88)90064-2

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

R Core Team. (2023). *R: A language and environment for statistical computing*. Foundation for Statistical Computing. https://www.R-project.org/

Rapp, D. N., & van Den Broek, P. (2005). Dynamic text comprehension: An integrative view of reading. *Current Directions in Psychological Science*, *14*(5), 276–279. https://doi.org/10.1111/j.0963-7214.2005.00380.x

Rapp, D. N., van den Broek, P., McMaster, K. L., Kendeou, P., & Espin, C. A. (2007). Higher-order comprehension processes in struggling readers: A perspective for research and intervention. *Scientific Studies of Reading*, *11*(4), 289–312. https://doi.org/10.1080/10888430701530417

Ray, M., & Magliano, J. P. (2015, November). *Automatic assessment of moment-to-moment comprehension processes* [Presented presentation]. Annual Meeting for the Society for Computers in Psychology, Chicago, IL.

Scardamalia, M., & Bereiter, C. (1992). Text-based and knowledge based questioning by children. *Cognition and Instruction*, *9*(3), 177–199. https://doi.org/10.1207/s1532690xci0903_1

Seifert, C. M., Robertson, S. P., & Black, J. B. (1985). Types of inferences generated during reading. *Journal of Memory and Language*, *24*(4), 405–422. https://doi.org/10.1016/0749-596X(85)90037-3

Seipel, B., Carlson, S. E., & Clinton, V. E. (2017). When do comprehender groups differ? A moment-by-moment analysis of think-aloud protocols of good and poor comprehenders. *Reading Psychology*, *38*(1), 39–70. https://doi.org/10.1080/02702711.2016.1216489

Singer, M. (1988). Inferences in reading comprehension. In M. Daneman, G. E. Mackinnon, & T. G. Waller (Eds.), *Reading research: Advances in theory and practice* (Vol. 6, pp. 177–219). Academic Press.

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

Talwar, A., Tighe, E. L., & Greenberg, D. (2018). Augmenting the simple view of reading for struggling adult readers: A unique role for background knowledge. *Scientific Studies of Reading*, *22*(5), 351–366. https://doi.org/10.1080/10888438.2018.1450410

Trabasso, T., & Magliano, J. P. (1996). Conscious understanding during comprehension. *Discourse Processes*, *21*(3), 255–287. https://doi.org/10.1080/01638539609544959

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

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

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

van den Broek, P., & Helder, A. (2017). Cognitive processes in discourse comprehension: Passive processes, reader-initiated processes, and evolving mental representations. *Discourse Processes*, *54*(5–6), 360–372. http://dx.doi.org/10.1080/0163853X.2017.1306677

van den Broek, P., Lorch, R. F., Linderholm, T., & Gustafson, M. (2001). The effects of readers' goals on inference generation and memory for texts. *Memory & Cognition*, *29*(8), 1081–1087. https://doi.org/10.3758/BF03206376

van den Broek, P., Young, M., Tzeng, Y., & Linderholm, T. (1999). The Landscape model of reading: Inferences and the online construction of memory representation. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 71–98). Erlbaum.

VanLehn, K., Jones, R. M., & Chi, M. T. T. (1992). A model of the self-explanation effect. *Journal of the Learning Sciences*, *2*(1), 1–59. https://doi.org/10.1207/s15327809jls0201_1

Wang, Y.-H. (2016). Reading strategy use and comprehension performance of more successfuland less successful readers: A think-aloud study. *Educational Sciences: Theory & Practice*, *16*(5), 1789–1813.

Whitney, P., Ritchie, B. G., & Clark, M. B. (1991). Working-memory capacity and the use of elaborative inferences in text comprehension. *Discourse Processes*, *14*(2), 133–145. https://doi.org/10.1080/01638539109544779

Williamson, K., & Kizilcec, R. (2021). Effects of algorithmic transparency in Bayesian knowledge tracing on trust and perceived accuracy. In A. N. Rafferty, J. Whitehill, C. Romero, & V. Cavalli-Sforza (Eds.), *Proceedings of The 13th International Conference on Educational Data Mining (EDM '20)* (pp. 338–344). Virtual Conference. https://educationaldatamining.org/edm2021

Appendix A. Self-explanations of independent and combined strategies for the cell division text

Strategy	Sent No.	Target Sentence	Constructed Response
None	1	In the first stage, mitosis, one complete set of chromosomes goes to each daughter cell.	*In the first stage, mitosis completes set of chromosomes goes to each daughter cell.*
	16	This results in cells with several nuclei.	*Yes, I think so.*
Paraphrase only	3	Each daughter cell will have about half of the organelles from the cytoplasm of the parent cell.	*The daughter caries half of each parents' organelles*
	8	In the second phase of mitosis, called metaphase, the chromatids become aligned at the midregion, or equator, of the cell.	*This shows that the metaphase is when the chromatids are align at the middle of the cell.*
Bridging only	2	This process is cytokinesis.	*After all of the genetic information of the parent cell is transferred to the daughter cell, then the cytoplasm and its content divides.*
	14	Therefore, before cytokinesis, there is growth in the size of the cytoplasm and in the number of its organelles.	*This sentence states that many organelles are formed from components in the cytoplasm so when the cell divides during cytokinesis, it provides opportunity for more to be formed.*
Elaboration only	6	These filaments resemble a flower and are called asters.	*These structures must allow for communication and connection between the cells.*
	8	In the second phase of mitosis, called metaphase, the chromatids become aligned at the midregion, or equator, of the cell.	*This stage can be observed under a microscope and is sometimes the easiest to differentiate from other stages. This is the final stage before external changes can be seen on the cell membrane.*
Para-bridging	6	These filaments resemble a flower and are called asters.	*The sentence is saying that the filaments of an animal cell look like a flower and are called asters. This means that the animal cell can be easily identified because one must look for the flower shaped asters.*
	10	This causes the chromatids to separate into daughter chromosomes.	*The genetic information is finally separated as the cell begins to divide*
Para-elaboration	1	In the first stage, mitosis, one complete set of chromosomes goes to each daughter cell.	*The daughter cell is the first to form in the first stage called mitosis. Each daughter cell recieves a complete set of chomosomes. When a complete set of chromosomes is not fully administered to each daughter cell, downs-syndrome or other mental handicaps might occur.*
	14	Therefore, before cytokinesis, there is growth in the size of the cytoplasm and in the number of its organelles.	*Since cytokinesis depends on certain variables, there is a growth in size of the cytoplasm and number of organelles to help the cytokinesis. It needs energy and other variables to continue, just as anything in an organism does.*
Bridge-elaboration	1	In the first stage, mitosis, one complete set of chromosomes goes to each daughter cell.	*When cells divide, the parent cell splits exactly into two creating an identical daughter cell.*
	9	Metaphase in human cells requires from two to six minutes.	*This process takes significantly longer than the first step of cell division within humans. This pattern could be completely different in other animals. All of this occurs every day when we don't even know it. Cells take turns going through this process.*
All	2	This process is cytokinesis.	*Cytokineses – Second stage. Cytokineses replicates the rest of the cell and it's contents like mitochondria and theother cell organs.*
	3	Each daughter cell will have about half of the organelles from the cytoplasm of the parent cell.	*Like humans cellshave organsthat allow the cells to complete their functionality. Each cell is given half of their needed organs from the parent cell. There is still another step to be done before each cell is complete.These cells are floating around in the goo that is the cytoplasm.*

The constructed response entries appear exactly as they were submitted. "Para-bridging" refers to the combination of paraphrasing and bridging. "Para-elaboration" refers to the combination of paraphrasing and elaboration. "Bridge-elaboration" refers to the combination of bridging and elaboration. "All" refers to the combination of paraphrasing, bridging, and elaboration.

Appendix B. Means and standard deviations of Self-Explanations (SEs) per coding category

	Overall		Dataset 1 – HD		Dataset 1 – RBC		Dataset 2 – CD		Dataset 3 – CD	
	n = 472; SEs = 5,858		*n* = 122; SEs = 1,098		*n* = 120; SEs = 1080		*n* = 77; SEs = 1,232		*n* = 153; SEs = 2,448	
Coding Category	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Copied or Retyped	**0.25**	0.43	**0.07**	0.25	**0.12**	0.33	**0.48**	0.50	**0.26**	0.44
Partially Copied	**0.21**	0.40	**0.05**	0.23	**0.09**	0.28	**0.41**	0.49	**0.22**	0.42
Entirely Copied	**0.04**	0.20	**0.01**	0.12	**0.04**	0.18	**0.07**	0.26	**0.04**	0.20
Evaluative Statements	**0.01**	0.10	**0.01**	0.11	**0.01**	0.07	**0.02**	0.12	**0.01**	0.09
Misconceptions	**0.02**	0.14	**0.02**	0.14	**0.01**	0.09	**0.09**	0.09	**0.03**	0.17
Comprehension	**0.02**	0.12	**0.03**	0.16	**0.03**	0.18	**0.01**	0.06	**0.01**	0.09
Paraphrase Presence	**0.91**	0.29	**0.94**	0.23	**0.95**	0.21	**0.83**	0.37	**0.91**	0.28
Lexical Change	**0.59**	0.49	**0.63**	0.48	**0.57**	0.50	**0.52**	0.50	**0.61**	0.49
Syntactic Change	**0.52**	0.50	**0.48**	0.50	**0.46**	0.50	**0.38**	0.49	**0.64**	0.48
Bridging Presence	**0.60**	0.49	**0.52**	0.49	**0.57**	0.49	**0.65**	0.48	**0.62**	0.48
Local Bridge	**0.26**	0.44	**0.23**	0.42	**0.29**	0.45	**0.28**	0.45	**0.26**	0.44
Distal Bridge	**0.34**	0.47	**0.29**	0.46	**0.28**	0.45	**0.38**	0.48	**0.37**	0.48
Elaboration Presence	**0.06**	0.24	**0.08**	0.28	**0.07**	0.26	**0.05**	0.23	**0.05**	0.22
Overall Quality Mean	**1.57**	0.78	**1.78**	0.80	**1.83**	0.82	**1.41**	0.78	**1.43**	0.69
Poor	**0.06**	0.23	**0.02**	0.15	**0.04**	0.19	**0.11**	0.31	**0.05**	0.23
Fair	**0.44**	0.50	**0.38**	0.48	**0.33**	0.47	**0.44**	0.50	**0.53**	0.50
Good	**0.38**	0.48	**0.39**	0.49	**0.40**	0.49	**0.38**	0.48	**0.36**	0.48
Great	**0.12**	0.33	**0.21**	0.41	**0.23**	0.42	**0.07**	0.26	**0.06**	0.24

"SEs" refers to Self-Explanations. HD refers to Heart Disease; RBC refers to Red Blood Cells; CD refers to Cell Division.