



Received: 21 December 2017
Accepted: 21 December 2018
First Published: 17 January 2019

*Corresponding author: Christian Soto, Department of Spanish, University of Concepción, Victor Lamas 1290, Concepción 4030000, Chile
E-mail: christiansoto@udec.cl

Reviewing editor:
Richard Kruk, Psychology, University of Manitoba, Canada

Additional information is available at the end of the article

EDUCATIONAL PSYCHOLOGY & COUNSELLING | RESEARCH ARTICLE

Reading comprehension and metacognition: The importance of inferential skills

Christian Soto^{1*}, Antonio P. Gutiérrez de Blume², Mathew Jacovina³, Danielle McNamara³, Nicholas Benson⁴ and Bernardo Riffo¹

Abstract: We explored relations between reading comprehension performance and self-reported components of metacognition in middle-school children. Students' self-reported metacognitive strategies in planning and evaluation accounted for significant variance in reading comprehension performance on questions involving inferences. In Study 2, middle school students read a science text then made predictions about how they would perform on a comprehension test. Students' *metacomprehension accuracy* was related to their performance at different levels of understanding. Students' text-based question performance accounted for significant variance in metacomprehension accuracy for text-based questions, and inference-based question performance accounted for significant variance in metacomprehension accuracy for inference-based questions. Results from the two studies suggest that metacognitive and metacomprehension knowledge is aligned with the level of information given in text, and is related to deeper understanding of texts, particularly for inferential information. We discuss the implications of these



Christian Soto

ABOUT THE AUTHOR

Christian Soto Our research team is currently investigating metacomprehension in reading and how this is related to other metrics of metacognitive monitoring, most notably learners' ability to accurately report what they know or do not know about a topic. We are also examining whether intelligent tutoring systems based on artificial intelligence can effectively and efficiently train reading comprehension skills in children and adolescents. We believe these lines of inquiry are essential to inform not only educational practice of teachers in classrooms but educational policy as well such as funding decisions for school systems and educational research.

PUBLIC INTEREST STATEMENT

Understanding how adolescents learn is an important endeavor for learning scientists. We explored relations between reading comprehension performance and self-reported components of metacognition in middle-school children. In Study 1, students' self-reported metacognitive strategies in planning and evaluation significantly positively related to reading comprehension performance on questions involving inferences. In Study 2, middle school students read a science text then made predictions about how they would perform on a reading comprehension test. Students' reading comprehension accuracy was related to their performance at different levels of the text. Students' text-based question performance, which relies on superficial understanding of texts, was positively related to reading comprehension accuracy for text-based questions, and inference-based question performance, which requires a deeper understanding of the text because it necessitates linking what one reads with prior knowledge of the topic, was significantly positively related to reading comprehension accuracy for inference-based questions. This information will help classroom teachers to tailor reading comprehension interventions to student needs to encourage deeper understanding of texts.

findings and how future research on absolute metacomprehension accuracy should consider different levels of understanding.

Subjects: Psychological Science; Educational Research; Educational Psychology

Keywords: reading comprehension; metacomprehension accuracy; inferential skills; metacognition

1. Introduction

Metacognition is a complex construct that has a rich history in the research literature. Flavell (1979) defined “metacognition” as taking one’s own cognition as the object of thought. Since Flavell’s seminal work on metacognition, researchers have strived to provide more nuanced definitions of this complex concept. Nelson and Narens (1990, 1994), for instance, further conceptualized metacognition as comprising two fundamental processes, monitoring and control. Subsequent research has more clearly specified metacognition as inclusive of two main dimensions: knowledge of cognition and regulation of cognition, which subsume more fine-grained sub-processes (Schraw & Dennison, 1994). Knowledge of cognition involves declarative knowledge (repertoire of strategies to employ during learning), procedural knowledge (steps necessary to employ strategies) and conditional knowledge (the knowledge of where, when, and why to employ more versus less successful strategies, given task demands). Regulation of cognition, on the other hand, incorporates those processes necessary to monitor and control learning: planning, information management strategies, debugging strategies, evaluation of learning, and comprehension monitoring (Schraw & Dennison, 1994). We focus on one of the sub-processes of regulation in the present study, comprehension (metacognitive) monitoring.

Comprehension monitoring involves the skill to accurately, efficiently, and effectively monitor the learning task and control subsequent actions to successfully achieve learning objectives (Schraw & Dennison, 1994). Monitoring and control are hypothesized to be cyclical, reciprocal processes in this context (Nelson & Narens, 1990, 1994). Throughout time, various indices and metrics of metacognitive monitoring have been proposed. For the purposes of this study, metacognitive monitoring accuracy has been defined as a feeling-of-knowing (FOK) judgment in which learners make judgments about future performance on a task (e.g., test or exam, known as prospective judgments), albeit it is also possible to assess past performance (retrospective) (Schraw, 2009). The choice of whether to measure global (a holistic, overall) or local (item-by-item) judgment also has implications for the interpretation of metacognitive monitoring (Schraw, 2009). The extent to which individuals’ judgments of their performance are congruent with actual performance is known as monitoring accuracy (henceforth in our research, metacomprehension accuracy) whereas the mismatch between judgments and performance are known as metacomprehension bias or error (Boekaerts & Rozendaal, 2010; Efklides, 2008; Winne & Nesbit, 2009; i.e., overconfidence and underconfidence). Metacomprehension accuracy, as a metric of metacognitive monitoring, has been studied using absolute versus relative judgments and with ease-of-learning judgments and judgments of learning (see Schraw, 2009, for a detailed review of these metacognitive monitoring indices).

Metacognitive monitoring has been employed in many domains. However, we focused on reading for the present studies. Reading comprehension is the set of skills that the subjects invoke to generate a mental representation of the text that is sufficiently coherent and rich enough to adequately understand the material that is being read. This presupposes the implementation of diverse cognitive tasks that go from the understanding of the words, relations between sentences and paragraphs, as well as capturing the sense of the text as a whole, among others. As theory has shown, students are able to process the content of the text at least in two levels of depth, base (henceforth, text-based) and inferential (Kintsch, 1998), depending on how the reader relates the ideas of the text, eventually incorporating previous knowledge and re-elaboration of the ideas

proposed in the written material. In the case of reading, metacomprehension involves metacognitive processes that support comprehension, allowing readers to evaluate the understanding in progress, and tentatively implement the necessary adjustments to improve the level of coherence of the mental representations generated in the reading process. Thus, in this study, we explored how global reading metacomprehension absolute accuracy, self-reported use of reading strategies, and reading performance are associated.

Efficient consumers of information possess an awareness of what they know as well as what they do not know and understand the specific actions they need to take to maximize their learning efficiency. This awareness is referred to as *metacognitive knowledge* and has been identified as an important part of successful learning (Baker & Beall, 2009; Palincsar & Brown, 1987; Schraw & Dennison, 1994). The ability to monitor the learning process is a key component of metacognition that allows people to recognize when their understanding of incoming information is not meeting their standards (Boekaerts & Corno, 2005; Flavell, 1979; Nelson & Narens, 1990, 1994; Zimmerman, 2000). Once people recognize a deficiency in their understanding they are better able to regulate their behavior and improve their understanding (Gutierrez & Schraw, 2015; Hacker, Bol, & Bahbahani, 2008; Schraw, 1998). Successful readers, for example, must determine when they have acquired sufficient knowledge from a text. If they recognize that they have not reached an acceptable level of understanding, they engage in additional processing, both processes (i.e., monitoring and control) which are essential in reading metacomprehension.

Effective metacognitive strategies involve more than simply making it to the end of a text and assuming an acceptable level of understanding of the text. When readers have an accurate understanding of their knowledge about a text, their metacomprehension accuracy is said to be high. In cases of poor metacomprehension accuracy, however, readers are less able to appropriately regulate their efforts. For example, consider a conscientious student studying for a biology midterm. Although she may spend ample time studying, if her metacomprehension is low, she may not know when to stop covering a certain topic because she cannot accurately gauge her level of understanding. Similarly, she would be less able to choose particular topics on which to focus, again because she is unsure just how well she comprehends each topic. Alternatively, she may be overly sure or confident of herself, and thus, yield the same outcome of poor metacomprehension. Indeed, poor metacomprehension has been linked to lower achievement (Bol & Hacker, 2001; Hacker et al., 2008; Zabrocky, Moore, Agler, & Cummings, 2015) and less well-regulated study (Thiede, 1999).

Unfortunately, students' metacomprehension accuracy is generally not proficient (Glenberg & Epstein, 1985; Glenberg, Sanocki, Epstein, & Morris, 1987; Griffin, Wiley, & Thiede, 2008). Primary goals in the field have thus far been to better understand metacomprehension and to develop techniques and interventions to enhance metacomprehension (e.g., Dunlosky & Rawson, 2005; Griffin et al., 2008; Gutierrez & Schraw, 2015). Despite many advances in the understanding of how metacognition influences reading comprehension, several questions remain. In the current project, we focus on two emerging topics regarding readers' metacomprehension. First, we examine which components of metacognitive knowledge are most predictive of reading comprehension performance for questions that require a text-based level understanding and for those that require a deeper level of understanding (i.e., requiring inferential reasoning). Second, we examine how reading comprehension skills relate to metacomprehension accuracy for both text-based and inferential understanding. In combination, these topics can elucidate the relations between metacognitive skills and reading comprehension, such as the ability to evaluate one's understanding and the resulting level of comprehension.

Despite empirical evidence supporting the benefits of metacognitive knowledge, its exact relation with reading comprehension is not entirely clear. For example, McNamara and Magliano (2009) claimed that the specific nature of the relation between metacognition and reading strategy use is unclear based on findings from a study using verbal protocols to relate

metacognition and self-explanations. Peronard (2005) found that an intervention aimed at increasing metacomprehension knowledge had little impact on reading comprehension. Similarly, Puente Jiménez and Alvarado (2009) found weak relations between a reading comprehension test (Batería de Evaluación de los Procesos Lectores en Secundaria; PROLEC-SE) and a reading awareness test (Escala de Conciencia Lectora; ESCOLA). Thus, students' explicit metacomprehension knowledge does not seem to be a consistently strong predictor of their reading comprehension. In the following sections, we will describe some reasons for these mixed findings, including concerns about the types of reading comprehension tests used in previous studies.

Before delving into the literature about how metacognition influences reading comprehension, it is important to disambiguate our conceptualization of comprehension monitoring. Researchers have described monitoring as including different processes. According to Hacker and colleagues, monitoring has often been discussed as including both the processes of evaluation and regulation (Hacker et al., 1994; Keener & Hacker, 2012). From this perspective, readers' monitoring would be said to be successful only when they, for example, both noticed that they did not understand some part of the text and deployed cognitive effort to remedy their understanding (e.g., by rereading the section). An alternative view of monitoring is that there is a clear distinction between the processes of monitoring (e.g., evaluating comprehension), and regulation (e.g., doing something to fix comprehension deficits) (Schraw & Dennison, 1994; Schraw & Moshman, 1995). We adopt this second view, as we believe it to be more useful when studying metacomprehension and its influence on reading comprehension.

1.1. Metacomprehension accuracy

As stated earlier, students' metacomprehension accuracy tends to be quite poor. Studies that find these discouraging results often ask students to read a text, make a prediction (or predictions) about how they will perform on a comprehension test, and then complete a comprehension test (e.g., Bol & Hacker, 2001; Bol, Hacker, O'Shea, & Allen, 2005; Dunlosky & Lipko, 2007; Hacker et al., 2008). Although many of the predictions are about students' overall comprehension level (e.g., asking how well students think they will do on an upcoming test), even when students are asked to predict performance regarding specific bits of information from a text (e.g., asking how well students think they will be able to recall a definition), accuracy is still quite low (Dunlosky, Rawson, & McDonald, 2002). The relation between students' predictions and their performance is considered a measure of monitoring; when students successfully evaluate their level of comprehension they should be quite accurate in their predictions.

Thiede and colleagues (Thiede, Griffin, Wiley, & Redford, 2009) conducted an extensive analysis of the relations between evaluation and reading comprehension. They found that the average correlation—in more than 40 studies—was about .27. Thiede et al. explained this low correlation by describing several concerns that could distort the relation between students' predictions of performance and their reading comprehension scores, or that could lead to low levels of precision in evaluation. The primary concern relates to a lack of evidence supporting the validity of scores from the reading comprehension tests used in previous studies. Weaver (1990) demonstrated that there are improvements of metacomprehension accuracy when several elements of the text are included in the comprehension assessment. Additionally, the assessment must include information from all sections of the text (Dunlosky, Griffin, Thiede, & Wiley, 2005). Therefore, the authors suggest that measures of reading comprehension must cover various components of the text as well as relations among these components. When a measure of comprehension focuses too heavily on certain material the correspondence between predictions and performance will, on average, be lowered. Dunlosky and Lipko (2007) also noted that metacomprehension accuracy can be influenced by the length of a text, such that for longer texts readers will have more trouble making accurate predictions of their comprehension, and hence, their metacomprehension accuracy will suffer (Thiede et al.,

2009). It is important to note, however, that they employed a relative accuracy index whereas we use absolute accuracy.

The other key explanation of poor metacomprehension accuracy is that readers may be using different cues to generate their judgments of understanding. Dunlosky et al. (2002) proposed the “levels of disruption” theory, which assumes that when readers make judgments about the understanding of a text they base this prediction on cues that are derived from disruptions that threaten the flow of reading. The *inference assumption* proposes that metacomprehension judgments are derived from inferences that people make based on instances of disruption. Many factors can cause disruptions, such as reading unfamiliar words, ambiguous pronouns, and so on. According to Dunlosky and his colleagues, as more disruptions occur people are increasingly likely to infer that the text has been misunderstood. The *accuracy assumption* proposes that metacomprehension judgments are a function of the degree to which key underlying judgments are predictive of comprehension test performance. As an example, readers might infer that the length of a text will influence their test performance and therefore will use this as a cue when making metacomprehension judgments. If length is correlated with the difficulty of a test then the accuracy of judgments will be relatively high; but this will not always be the case. Finally, the *representation assumption* proposes that disruptions can occur at different levels of representation of the text (i. e., text-based and situation model; Kintsch, 1998). When disruptions occur primarily at one specific level of readers’ metacomprehension, it is based on that level of the text. We propose that the representation assumption is especially important in metacomprehension and focus on this assumption throughout this paper.

1.2. Levels of representation and metacomprehension

According to the Construction-Integration Model (Kintsch, 1988, 1998), text comprehension involves multiple levels of information processing. The linguistics level involves recognizing words and understanding the syntactic links between them (Riffo, Reyes, Novoa, Véliz, & Castro, 2014). A second, text-based level involves the generation of meaning through the integration of propositions. A third level, known as the situation model, integrates textual information with additional information from the reader’s prior knowledge (Kintsch & Rawson, 2007). Generating inferences is perhaps the most important mental operation in the construction of the situation model. Inferential processes result in readers generating information that is not directly stated in the text (i.e., by invoking prior knowledge). Importantly, this new information yields deeper understanding because readers include these semantic elements to their mental representation of text information. Thus, inferences play an important role in the quality of mental representations (Vieiro & Gómez, 2004).

Many researchers have found that higher inferential abilities are linked to reading comprehension. For example, research has shown that children who experience comprehension difficulties often lack skills needed to generate inferences (Cain, Oakhill, Barnes, & Bryant, 2001). Similarly, high inferential abilities such as those that require individuals to use prior knowledge and information within-text to reach conclusions are associated with good reading comprehension (Cain & Oakhill, 1999). Other lines of work also indicate that teaching inferential skills can improve reading comprehension. McNamara, for example, developed Self-Explanation Strategy Training (SERT), the goal of which is to teach students reading strategies that involve self-explanation and encourage the generation of inferences during reading (McNamara, 2004, 2017). An evaluation of SERT showed that students who are low in prior knowledge benefitted from SERT (McNamara, 2004). This same study also found a positive relation between the generation of inferences that establish connections between text elements (i.e., bridging inferences) and performance with reading comprehension questions.

Prior research, thus, supports the idea that inferencing is an important factor influencing the quality of readers’ mental representations. Inferential skills should, therefore, influence readers’ metacomprehension. Specifically, we argue that metacomprehension judgments,

which require careful evaluation, should be influenced by readers' inferential skills. Because the assumption of representation posits that readers will base metacognitive judgments on disruptions at particular levels of representations, readers who have higher or lower inferential skills will base their judgments on different information. For example, when readers with higher inferential skill read a text they will be more likely to encounter disruptions at the situation model level of representation. In other words, because they are frequently generating inferences and tying information to their prior knowledge they will be more likely to notice discrepancies at this deeper level of understanding. Conversely, readers with lower inferential skill will generate fewer inferences and be less likely to notice disruptions at the situation model level—instead noticing more difficulties in building their text-based level knowledge.

Other researchers have argued that the quality of readers' mental representations relate to metacomprehension accuracy. For example, Thiede and colleagues (Thiede et al., 2009) noted that if readers are generating valid inferences, metacomprehension accuracy increases as they use cues based on the situation model rather than a surface level or text-based level (Dunlosky et al., 2005, 2002). This idea has been supported by some pioneering studies that have improved metacomprehension accuracy using techniques, during or after reading, that are intended to support more complete, readily accessible mental representations. These techniques have included summarization or keywords after a delay, self-explanation while reading, and the construction of concept maps. To relate this to Kintsch's model, if readers write a delayed summary after completing a text they will likely have access to a more complete situation model of the text. Similarly, when readers generate a list of keywords for a text, their metacomprehension judgments more closely align with comprehension test questions, allowing students to make more valid judgments. For example, in a study by Anderson and Thiede (2003), metacomprehension accuracy was dramatically higher for a group that made a delayed summary ($r = .60$) than for the control group ($r = .26$); this result was replicated for both longer and shorter texts. Again, these results suggest that metacomprehension is influenced by readers' mental representations and that encouraging readers to engage in inferential processes can improve metacomprehension accuracy.

1.3. Purpose of the studies

Different approaches can be used to investigate how readers' metacognitive knowledge links to their level of comprehension at multiple levels (i.e., linguistic, text-based, and situation levels). To this end, we have divided our research into two studies.

1.4. Study 1

In Study 1, students completed a self-report survey about their metacognitive knowledge, specifically as it relates to reading. Scores on three subcomponents of metacognition were measured, including planning (i.e., engaging in processes aimed at preparing for a reading task), monitoring (i.e., recognizing problems with comprehension and making necessary adjustments during reading), and evaluation (i.e., assessing and recognizing successes and failures during reading). The main research questions that guided Study 1 were:

- (1) To what extent are self-reported planning, monitoring, and evaluation related to performance on text-based and inferential comprehension questions among a cohort of Chilean middle school students?
- (2) How much variance do planning, monitoring, and evaluation account for in performance regarding text-based and inferential comprehension questions?

H₁: Based on our review of the literature, we predicted that metacognitive knowledge will be more highly associated with students' performance on comprehension questions that require inferencing (versus text-based).

H₂: In line with our previous hypothesis, we expected that planning, monitoring, and evaluation would be better predictors of inferential comprehension questions than text-based questions.

1.5. Study 2

In Study 2, students made predictions about their upcoming performance on a comprehension test for a text they just read. Metacomprehension absolute accuracy was then calculated for performance on text-based and inference level questions. The following research questions guided Study 2:

- (3) To what extent does reading comprehension performance predict metacomprehension absolute accuracy?
- (4) Does the relation between metacomprehension absolute accuracy and reading comprehension performance vary by question type (i.e., inferential, text-based)?

H₃: We expected higher performance on reading comprehension to significantly positively predict metacomprehension accuracy.

H₄: We expected reading comprehension performance to be more highly positively correlated for inferential level than text-based level judgments. Moreover, we expected reading comprehension and metacomprehension absolute accuracy to be greater for text-based level judgments compared to inferential level judgments.

2. Study 1

2.1. Method

2.1.1. Participants

A sample of 190 7th and 8th grade students were recruited from four different Chilean schools (two public schools and two private schools). Their ages ranged from 11–13 years of age ($M = 12.25$, $SD = 1.17$). The schools were public and private and had similar performance on the National System of Evaluation of the Quality of Education, which was in turn similar to the national average. Entire randomly chosen classes participated in this study. There were 93 females and 97 males. The 190 students were distributed as follows: 99 students from the public education system and 91 from the private education system. The students belonged to two courses of different schools of both educational systems (private and public). For this reason, the final sample was distributed as follows: Public School 1: 23 7th grade students and 25 8th grade students; Public School 2: 20 7th grade students and 31 8th grade students; Private School 1: 23 7th grade students and 28 8th grade students; and Private School 2: 21 7th grade students and 19 8th grade students.

2.1.2. Measures

2.1.2.1. *Metacomprehension skills. Escala de Conciencia Lectora* (ESCOLA; Reading Awareness Scale) is a reading awareness scale that measures metacomprehension skills in Spanish-speaking students between 8 and 13 years of age. It consists of 56 multiple-choice questions with three possible answers for each. ESCOLA assesses three different dimensions of metacomprehension: Planning, Monitoring, and Evaluation. Each question has a correct answer, and an answer that gives partial credit. Thus, the measure is designed to capture students' metacognitive competency by allowing them the opportunity to explain their answers and engage in self-explaining. Planning questions measure readers' knowledge of how to select appropriate reading strategies aimed at achieving reading goals. Monitoring questions measure students' ability to adjust attention and effort during the reading task. Evaluation questions measure students' understanding about whether appropriate understanding was achieved. Table 1 presents sample questions (translated by one of the authors of this paper) that exemplify each dimension. Following Hacker's definition, which we presented previously, ESCOLA's Monitoring dimension is closer to the concept of

Table 1. Sample questions from the three ESCOLA dimensions

PLANNING
Before you start reading, what do you do to help in the reading process?
a) I do not make any plans, I just start reading. [0 points]
b) I consider why I'm going to read. [2 points]
c) I choose a comfortable place to read. [1 point]
MONITORING
If you are reading a book and find a paragraph difficult to understand, what do you do?
a) I stop to think about the problem and how to fix it. [2 points]
b) I do not keep reading because I cannot solve the problem. [0 points]
c) I continue to read to see if the meaning is clarified later. [1 point]
EVALUATION
In carrying out the activity of reading:
a) I think it is useful to assess whether I understood what was written. [2 points]
b) I think that evaluating understanding is good but that it should be done by an adult [1 point]
c) I do not think that evaluating understanding is helpful after finishing reading. [0 points]
Study 1.

regulation during comprehension, whereas the Evaluation dimension is closer to the self-assessment of comprehension. ESCOLA has been used and validated in Spanish-speaking populations in both Spain and Argentina, and it has demonstrated sound psychometric properties with respect to the inferences and conclusions one can draw from data collected from the instrument (Jiménez, Puente, Alvarado, & Arbillaga, 2009; Puente et al., 2009). The internal consistency reliability coefficients, Cronbach's α , for the ESCOLA scales in the current sample were acceptable ($\alpha = .69$, Planning; $.71$, Monitoring; $.73$, Evaluation).

2.1.2.2. Reading comprehension test. Students' reading comprehension was assessed using comprehension questions for a text about digestion that was created for this research, with collaboration from an expert in discourse processes. More specifically, the text wording and content were verified by Eileen Kintsch (Institute of Cognitive Science at the University of Colorado at Boulder) and the first author of this research, who used the Construction-Integration model to classify every question of the test prior to submitting it for independent expert review. Of the 20 test questions, 10 are basic text-based questions and 10 are inferential questions. In the case of the basic text-based questions, the information needed to answer the questions is explicitly available in the text while the information for the answers to the inferential questions must be inferred. For both types of questions, we used a rubric per question, with scores from 0 to 2, depending on how complete or partial the information included in the answer was. More complete, coherent, and cogent responses, for instance, received complete credit whereas those with incomplete or weak answers were given 0. Responses with sound arguments but with gaps or incomplete information were given partial credit (i.e., a score of 1). A similar evaluative tool was used by McNamara, Kintsch, Songer, and Kintsch (1996) on the theme heart disease to assess the impact of prior knowledge and cohesion on reading comprehension. This model (Kintsch, 1998) is the most relevant model of the comprehension in the field of discourse processing, and assumes different levels of mental representation of the text (linguistic, text-based and situation model). The text did not contain images and was 433 words long. The comprehension test consisted of 20 open-ended questions, half of which were text-based questions and half of which were inferential questions. Sample items for text-based and inferential questions can be found in the Appendix.

Students' responses were assigned a score from 0 to 2 points for each question, and hence, the range of possible scores was 0–40. Each student was thus given a total score, and scores corresponding to their text-based and inferential performance. The assumption behind the text-based/inferential distinction is that students who can represent a coherent mental model during or after the reading will be better able to answer inferential questions. Students who are only able to answer explicit questions (text-based) are limited in their understanding. Internal consistency reliability coefficients, Cronbach's α , for this measure were adequate: text-based = .74; inferential = .81.

2.1.3. Procedure

Institutional review board (IRB) approval was secured from all relevant institutions prior to any data collection activities. Students first completed the ESCOLA test, which lasted approximately 40 minutes. Immediately after, they were given an expository text regarding the endocrine and gastro-intestinal systems and digestion. Students were given 50 minutes to read the text, after which the text was collected and the reading comprehension test was administered.

2.1.4. Data analysis

We screened the data for outliers within the scales of ESCOLA as well as the reading comprehension measure and evaluated the data for requisite assumptions prior to data analysis. The outlier analysis revealed the presence of 16 outliers (6 within planning and 10 within evaluation of the ESCOLA subscales). These outliers were detected using the casewise diagnostic command in regression by specifying standardized residuals beyond three standard deviations. Neglecting to address outliers may undermine the trustworthiness of results due to the undue influence these outliers exert on variable descriptive statistics (Tabachnick & Fidell, 2013). Thus, these outliers were omitted from data analyses, leaving 174 complete cases for analysis. These data met all assumptions, including normality, homoscedasticity, and linearity. Therefore, data analysis proceeded without making any adjustments to the data. Descriptive statistics for the ESCOLA scales and reading comprehension performance are presented in Table 2. Pearson's zero-order correlation coefficients were calculated to answer part of the first research question, and these are presented in Table 3. We subsequently conducted a series of simultaneous/standard ordinary least squares regressions to answer the second research question, on proportion of variance in comprehension performance accounted for by each component of metacomprehension, adjusting the a priori p -value using the Bonferroni adjustment for multiple analyses.

2.2. Results

Correlation coefficients in Table 3 show that all correlations were in the theoretically expected direction (i.e., positive). Interestingly, the planning and evaluation metacomprehension components correlated more strongly with inferential comprehension questions than with text-based comprehension performance, except for monitoring, which correlated more strongly with text-based comprehension questions. Results of the simultaneous regression with text-based comprehension questions as the criterion revealed that planning, monitoring, and evaluation were not significant predictors of text-based comprehension question performance, $F_{(3,170)} = 2.45$, $p = .06$. As would be expected based on Table 4, the closest predictor of performance on text-based comprehension to reaching statistical significance was monitoring ($p = .07$).

The simultaneous regression results with inferential comprehension questions as the criterion were statistically significant, $F_{(3,170)} = 7.38$, $p = .001$, $R^2 = .12$. However, only planning and evaluation were significant predictors, albeit evaluation was the strongest predictor of inferential comprehension question performance (see Table 4). Specifically, questions that related to students' evaluation of themselves as readers and their ability to plan effectively were related to their inferential performance.

Table 2. Descriptive statistics of ESCOLA scales and performance on inferential and text-based reading comprehension

Variable	M	SD	Minimum	Maximum	Skew	Kurtosis
ESCOLA						
Planning	37.31	4.82	25.00	48.00	-0.19	-0.36
Monitoring	22.99	3.31	15.00	30.00	-0.35	-0.42
Evaluation	19.69	2.67	13.00	26.00	-0.23	-0.26
Reading Comprehension						
Text-Based	7.82	3.10	1.00	12.00	-0.34	-1.06
Inferential	10.88	4.29	0.00	19.00	-0.36	-0.45

Study 1.
 N = 174

Table 3. Zero-order correlation matrix between the escola performance items and planning, monitoring, and evaluation

Variable	1	2	3	4	5
1. Text-based	-	.62**	.13*	.17*	.15*
2. Inferential		-	.22*	.10	.31**
3. Planning			-	.42**	.33**
4. Monitoring				-	.24*
5. Evaluation					-
Skew	-0.38	-0.35	-0.47	-0.77	-0.49
Kurtosis	-1.07	-0.44	-0.04	0.32	0.46

Study 1.
 N = 174
 * $p < .05$
 ** $p < .01$ (one-tailed)

Table 4. Standard regression results with text-based and inferential reading comprehension as outcomes and planning, monitoring, and evaluation as predictors

Predictor	B^+ (CI _{95%})	β^-	t	p
Text-Based Performance				
Planning	0.03 (-0.08, 0.14)	0.05	0.73	0.46 ^{ns}
Monitoring	0.12 (-0.04, 0.27)	0.12	1.48	0.14 ^{ns}
Evaluation	0.12 (-0.06, 0.30)	0.10	1.26	0.21 ^{ns}
Inferential Performance				
Planning	0.13 (0.02, 0.27)	0.15	2.02	0.039*
Monitoring	-0.04 (-0.24, 0.17)	-0.03	-0.34	0.73 ^{ns}
Evaluation	0.44 (0.19, 0.68)	0.27	3.53	0.001**

Study 1.
 N = 174 * $p < .05$ ** $p < .01$ ^{ns}Non-significant
 B^+ = Unstandardized regression coefficients and their 95% confidence interval (CI_{95%})
 β^- = Standardized regression coefficients

3. Study 2

3.1. Method

3.1.1. Participants

A distinct sample of 87 7th and 8th grade students—an independent sample from Study 1—were recruited from four different Chilean schools (two public schools and two private schools). As in Study 1, these schools had scores similar to the national average on the National System of Evaluation of the Quality of Education. Again, students from entire, randomly chosen classes participated in the study. This led to a similar distribution of males and females. Their mean age was 12.75 ($SD = 0.76$).

3.1.2. Measures

3.1.2.1. Reading comprehension test. Students' reading comprehension was assessed using comprehension questions for two texts: one about extinction and one about the endocrine system (Thiede, Wiley, & Griffin, 2011). These texts have been used in studies seeking to assess metacomprehension accuracy (e.g., Thiede et al., 2011). Each of the two text is an average length of 380 words and includes five multiple-choice text-based questions and five multiple-choice inferential questions, for a total of 20 questions. Each question had one correct answer, and hence, the total possible score was 10. We chose this approach because we did not want to merely replicate texts from Study 1 but to remain consistent with the procedure employed by Thiede et al. (2011). The Kuder-Richardson 20 internal consistency reliability coefficients for the tests were as follows: endocrine = .72; extinction = .69.

3.1.2.2. Metacomprehension accuracy. Immediately after reading each text, students made one holistic, overall judgment regarding how many questions out of five they believed they would answer correctly on an upcoming test; this procedure was done for text-based and inferential questions separately. In the metacognitive monitoring literature, these FOKs of future performance are known as *global prospective confidence* in performance judgments. It is noteworthy to alert the reader that this single global judgment is a course-grained analysis; including local judgments (i.e., item-by-item) would have provided a finer-grained analysis. During analyses, these judgments were compared with students' actual performance on text-based and inferential questions separately, thereby yielding an *absolute monitoring accuracy index*, one each for text-based and inferential questions. Some studies calculate metacomprehension accuracy by calculating the correlation between judgments and performance (choosing a relative accuracy measure; Schraw, 2009). However, we elected to use the *absolute accuracy* approach. Specifically, scores were calculated by comparing participants' global predictions of performance judgments (again, immediately after reading the text but before the test of performance on the text, or prospective) against their actual performance scores for each set of five text-based and five inferential questions respectively. More specifically, we subtracted their average global prospective judgment scores from their actual performance scores (i.e., number of correct responses on the text-based and on the inferential questions). Comparing predictions of performance against actual performance yielded continuous, absolute metacomprehension accuracy scores (henceforth, "metacomprehension accuracy"), as described by Schraw (2009); one text-based metacomprehension accuracy score, and one inference-based metacomprehension accuracy score for each passage. A score of "0" indicates perfect accuracy; on the other hand, the higher the value, and thus the farther away from "0", the greater the metacomprehension inaccuracy. Thus, the higher the score, the less accurate the participant's predictions of comprehension performance on the text.

We elected to use this measure because measures of absolute accuracy can be used to gauge metacomprehension accuracy (Keener & Hacker, 2012). Students who are proficient at predicting their future performance with reading comprehension tasks are more likely to take appropriate actions that promote metacomprehension accuracy and learning (Gutierrez & Schraw, 2015; Keener & Hacker, 2012; Schraw, Potenza, & Nebelsick-Gullett, 1993), and thus,

are of interest to the current study. When students are well calibrated, they are better able to learn from their studied materials (Winne, 2004) because they can take appropriate actions when necessary (Hacker et al., 2008), given task demands.

3.1.3. Procedure

Students completed the reading task for the first text, made a prediction about their subsequent performance, and then answered comprehension questions corresponding to that text. They then repeated this process for the second text. In total, the study lasted 60 minutes. The counterbalance for this second study was implemented by presenting the texts alternately in schools of the same type. More specifically, in 7th grade of one of the public schools, students were first asked to read and make performance judgments about the text that deals with extinction and then about the text that deals with the endocrine system. In the other public school, the order was reversed; that is, first we worked with the text on the endocrine system and then on the text about extinction. In the case of the two private schools, the same procedure for counterbalancing texts was applied.

3.1.4. Data analysis

Data screening and assumption testing procedures were conducted prior to data analysis. Data met the assumptions of normality and linearity. Further, an outlier analysis detected no cases that would be deemed outliers, and hence, all cases were retained for analysis. We subsequently conducted a series of standard/simultaneous ordinary least squares regressions, with each metacomprehension accuracy score—inferential and text-based extinction as well as inferential and text-based endocrine—serving as the criterion in each model respectively. Finally, in order to evaluate the effects of question level type (text-based, inferential) on reading comprehension performance and metacomprehension accuracy, we conducted two separate one-way multivariate analyses of variance (MANOVAs), one for endocrine and extinction comprehension performance as outcomes and the other for metacomprehension accuracy scores for endocrine and extinction questions as outcomes. We conducted two separate MANOVAs because we sought to more clearly disentangle the unique effects of performance and metacomprehension accuracy. We controlled for Type I error rate inflation using the Bonferroni adjustment.

3.2. Results

Table 5 presents the descriptive statistics of performance and Table 6 presents metacomprehension accuracy scores as a function of the text type. Table 7 contains the Pearson's zero-order correlation coefficients. Negative correlations between performance and metacomprehension accuracy can be interpreted as better comprehension performance being related to better metacomprehension accuracy, as the higher the comprehension performance, the lower the miscalibration. This was a function of the method we chose to compute absolute metacomprehension accuracy.

Correlation coefficients in Table 7 show that metacomprehension accuracy was strongly related to question type (i.e., inferential question performance more strongly correlated with metacomprehension accuracy on those items and text-based performance correlated more strongly with text-based metacomprehension accuracy scores). Results of a standard regression revealed that only inferential question performance on the extinction test was a significant predictor of inferential question metacomprehension accuracy, $F_{(4,82)} = 42.21$, $p = .001$, $R^2 = .53$. Similarly, text-based question performance on the extinction test predicted text-based question metacomprehension accuracy, $F_{(4,82)} = 24.21$, $p = .001$, $R^2 = .37$. The same pattern held for the endocrine text. Inferential question performance on the endocrine test predicted inferential question metacomprehension accuracy, $F_{(4,82)} = 29.66$, $p = .001$, $R^2 = .40$; and text-based question performance on the endocrine test was the only significant predictor of text-based question metacomprehension accuracy, $F_{(4,82)} = 28.34$, $p = .001$, $R^2 = .40$. Table 8 presents the results of the standard regression models.

Table 5. Descriptive statistics of performance by question type and type of test

Question Type	Endocrine				Extinction			
	M	SD	Skew	Kurtosis	M	SD	Skew	Kurtosis
Inferential	1.67	1.01	0.54	0.29	1.81	1.16	0.26	0.02
Text-based	2.43	1.32	0.08	-0.81	2.98	1.53	-0.28	-1.04

Study 2.
 N = 87

Table 6. Descriptive statistics for metacomprehension accuracy by question type and type of test

Question Type	Endocrine				Extinction			
	M	SD	Skew	Kurtosis	M	SD	Skew	Kurtosis
Inferential	2.13	1.13	0.41	0.16	2.06	1.15	0.14	-0.54
Text-based	1.57	1.12	0.35	-0.51	1.52	1.04	0.96	1.03

Study 2. Higher values indicate lower metacomprehension accuracy.
 N = 87

Table 7. Correlation matrix of reading comprehension performance and metacomprehension accuracy for endocrine and extinction questions by text-based- and inferential-level questions

Variable	1	2	3	4
1. Endocrine Performance	-	.18*	-.63**	.01
2. Extinction Performance	.92**	-	-.09	-.60**
3. Endocrine	-.64**	-.52**	-	-.09
4. Extinction	-.63**	-.73**	.33**	-

Study 2. Correlation coefficients above the diagonal are for text-based level questions and those below the diagonal are for inferential-level questions.
 N = 87

These analyses suggest that students' appraisal of their understanding will be somewhat dependent on whether they understood the text at a deep or surface level. Because students were not asked to make separate predictions for their inferential and text-based comprehension performance (students, in fact, were not told anything about these distinctions), these different patterns of relations seem to emerge without prompting or priming.

With respect to our final research question, correlations between reading comprehension performance and metacomprehension accuracy revealed that correlation coefficients were higher in absolute value for inferential-level questions (Pearson's correlations ranged from $r = .32$ to $r = .92$ in absolute value) than for text-based-level questions (Pearson's correlations ranged from $r = .08$ to $r = .63$ in absolute value; see Table 7 for the correlation matrix). Because our absolute metacomprehension accuracy index is interpreted such that higher values indicate greater miscalibration, negative correlations between the comprehension performance and absolute metacomprehension accuracy measures suggest that the higher the comprehension performance the lower the miscalibration, and hence, the greater the metacomprehension accuracy.

Results of the one-way MANOVA with endocrine and extinction comprehension performance as outcomes revealed that question level (text-based, inferential) had a significant effect on the linear combination of dependent variables, multivariate $F_{(2,171)} = 17.33, p < .001, \eta^2 = .169$. Given these significant results, the univariate results were interpreted next. Question level had a significant effect on both sets of questions of the text: endocrine reading comprehension

Table 8. Standard regression results of text-based and inferential reading comprehension performance on absolute accuracy by endocrine and extinction topics

Predictor	B ⁺ (CI _{95%})	β ⁻	t	p
Text-Based Absolute Accuracy				
Extinction Performance				
Inferential	0.09 (-0.06, 0.25)	0.08	1.18	0.23 ^{ns}
Text-Based	-0.42 (-0.54, -0.30)	-0.61	-6.95	0.001**
Endocrine Performance				
Inferential	-0.08 (-0.25, 0.10)	-0.08	-0.91	0.36 ^{ns}
Text-Based	-0.51 (-0.66, -0.37)	-0.62	-7.17	0.001**
Inferential Absolute Accuracy				
Extinction Performance				
Inferential	-0.72 (-0.87, -0.57)	-0.73	-9.54	0.001**
Text-Based	0.01 (-0.10, 0.13)	0.13	0.17	0.86 ^{ns}
Endocrine Performance				
Inferential	-0.68 (-0.86, -0.50)	-0.65	-7.63	0.001**
Text-Based	0.03 (-0.11, 0.18)	0.04	0.41	0.67 ^{ns}

Study 2.

N = 87 *p < .05 **p < .01^{ns}Non-significant

B⁺ = Unstandardized regression coefficients and their 95% confidence interval (CI_{95%})

β⁻ = Standardized regression coefficients

performance, $F_{(1,172)} = 16.69, p < .001, \eta^2 = .088$; and extinction reading comprehension performance, $F_{(1,172)} = 32.18, p < .001, \eta^2 = .158$. Students performed significantly better in text-based level comprehension performance across both endocrine- ($M = 2.43, SD = 1.33$) and extinction-related questions ($M = 2.98, SD = 1.53$) when compared to inferential level questions (endocrine, $M = 1.68, SD = 1.07$; extinction, $M = 1.80, SD = 1.17$).

Findings regarding metacomprehension accuracy for the endocrine- and extinction-related passages indicated that question level (text-based, inferential) had a statistically and practically significant effect on overall metacomprehension accuracy for each passage, multivariate $F_{(2,171)} = 9.56, p < .001, \eta^2 = .101$. Univariate results revealed that question level had a significant effect on students' endocrine, $F_{(1,172)} = 10.91, p = .001, \eta^2 = .061$, and extinction, $F_{(1,172)} = 10.88, p = .001, \eta^2 = .060$, metacomprehension accuracy. Students exhibited significantly greater metacomprehension accuracy—for text-based (endocrine, $M = 1.57, SD = 1.12$; extinction, $M = 1.52, SD = 1.04$) than for inferential questions (endocrine, $M = 2.14, SD = 1.13$; extinction, $M = 2.07, SD = 1.16$).

In sum, findings of our two one-way MANOVAs indicate consistently that students tend to perform better and manifest greater metacomprehension accuracy for text-based level comprehension performance than for questions that require them to draw inferences based on information provided. This pattern generalized across both sets of questions of the text, as it held constant in both questions assessing students' knowledge of the endocrine system and those evaluating knowledge of factors related to extinction.

4. General discussion

In this series of studies, we sought to better understand how text-based and inferential performance related to self-reported measures of metacognition (Study 1) and students' metacognitive *global absolute* monitoring judgments of their own reading comprehension made immediately after the text was read (Study 2). Study 1 demonstrated a relation between students' evaluative reading knowledge, as measured by ESCOLA, and their performance on comprehension questions that required inferential reasoning. This suggests that being aware of self-reported use of

evaluative comprehension processes and how they relate to success in reading could be a key factor in supporting inferential comprehension. This is specifically related to declarative knowledge of reading strategies and the when, where, and why to apply strategies given task demands (i.e., conditional knowledge), which elucidates how those strategies influence comprehension (Jacobs & Paris, 1987; Palincsar & Brown, 1987; Schraw, 1994). Planning was also a significant positive predictor of inferential comprehension performance. This suggests that learners need more developed planning skills in order to more deeply comprehend texts, especially when those texts are complex, and thus, require inferential reasoning. Presumably, readers with proficient comprehension of texts incorporate planning as an important process to help them generate high-quality inferences. (Dunlosky et al., 2002; Schraw, 1994).

Study 2 demonstrated that students' global absolute metacomprehension accuracy was related to their performance at different levels of understanding, and that these relations were significantly higher for questions that required inferential reasoning than for declarative-related questions (i.e., text-based). Students who performed well on inferential questions tended to also have high global absolute metacomprehension accuracy for inferential questions. This implies that readers' evaluations of their comprehension may be dependent on their level of processing of a text (see Kintsch, 1988, 1998). This is consistent with the representation assumption from the levels of disruption theory (Dunlosky et al., 2002). Students may base their level of understanding on the disruptions they encounter at a particular level of processing; hence, if students are attempting to generate many inferences, they will likely use their success at generating these inferences to estimate their level of understanding. For readers who do not generate as many inferences (i.e., less skilled readers or perhaps readers who are exerting less effort while reading), estimating their level of understanding will be done at a different level. Thus, if the goal of metacomprehension is to accurately be aware of one's deeper level understanding, it will be necessary to have adequate inferential abilities. Because evaluation involves generating an appropriate judgment about the success of comprehension (e.g., Hacker, 2014; Shraw, 2009; Thiede et al., 2009), it makes theoretical sense that inferential skills are related. It is important to note that we employed global absolute monitoring judgments provided immediately after the text was read. Nevertheless, research shows that delayed judgments tend to be more accurate than immediate judgments (e.g., Van Overschelde & Nelson, 2006; Weaver & Kelemen, 1997).

Next, we propose an explanation of how inferential skills may influence metacomprehension and then provide suggestions for how this account can be supported and refuted. We base this both on the results of the current studies as well as a the levels of disruption theory (Dunlosky et al., 2002).

4.1. The influence of inferential processes on global absolute metacomprehension and understanding

The explanation about the different performance on metacomprehension accuracy is that the readers may be using different cues to generate their judgments of understanding. Dunlosky et al. (2002) proposed the "levels of disruption" theory, which assumes that when readers make judgments about the understanding of a text, they base this prediction on cues that are derived from three possible disruptions that threaten the flow of reading, inference assumption, accuracy assumption and representation assumption. The latter proposes that disruptions can occur at different levels of representation of the text. When disruptions occur primarily at one specific level, readers' metacomprehension accuracy is based on that level of the text rather than on high-quality inferences.

According to other research, readers who excel at text-based comprehension performance understand the information of the text depending on the explicit and local information between adjacent ideas (McNamara & Magliano, 2009; Soto, Rodriguez, & Gutierrez de Blume, 2018). However, these same readers have some limitations in the monitoring cues to generate a more accurate judgment about their comprehension because at times that explicit information involves different dimensions (e.g., explicit ideas, words or simply details of the text, etc.). In fact, Soto,

Rodriguez, and Gutierrez de Blume (2018) employed the error detection paradigm among students with intellectual disabilities and they found that internal inconsistencies were a significant predictor of metacomprehension accuracy. However, the researchers show that the inferential readers operate using more sophisticated cues like self-explanation or elaboration. Our assumption is that the mental representation of readers with high inferential comprehension performance involves more coherent representations of the text, which in turn, yields a greater alignment between performance judgments and actual performance (i.e., greater metacomprehension accuracy).

The fact that two different metacognitive profiles could affect how the reading processing is generated, and particularly which information is considered (or not) is useful information for the appropriate representation of the text. For example, McNamara (2004) concluded that the most cohesive texts benefit poor readers, yet they could pose an obstacle for proficient readers. This is especially interesting because it could be evidence that, although metacognition in general affects reading outcomes, the specific metacognitive cues that the students use are different depending on the reader's profile, as our findings indicate. In the same way, the results of our present study show that the inferential readers demonstrate less accurate monitoring about the explicit information and details (text-based level), and hence, they have difficulty accessing inferential representation (situation level model) of the text.

4.2. Implications for learning and instruction

Our studies indicate that students' understanding of text is a function of the type of material they are reviewing. This is important information for educators to have at their disposal because they can use it to explicitly teach and encourage students to engage with reading materials more deeply, and thus, better understand the information presented. This is especially important for complex information, such as concepts of biology used in Study 2—namely, extinction and the endocrine system—which oftentimes are difficult for students, especially middle school students, to fully grasp. Findings from this research suggest that inferential reasoning plays a crucial role in deeper levels of understanding of materials. Hence, educators should emphasize the importance of inferential reasoning in their instruction, and hence, assist students to improve not only comprehension but metacomprehension monitoring. Previous research has demonstrated that learners who are more proficient in their monitoring of learning are also more self-regulated and successful (e.g., Gutierrez & Schraw, 2015; Hacker et al., 2008; Nietfeld & Schraw, 2002; Schraw, 1994; Schraw & Dennison, 1994). Thus, our results have practical implications and application to classroom educators.

Although our current results cannot suggest which reading skills would be more important for novice readers to practice, they do suggest that interventions, such as the Interactive Strategy Training for Active Reading and Thinking (iSTART; McNamara, Levinstein, & Boonthum, 2004; Snow, Jacobina, Jackson, & McNamara, 2016), that teach metacognitive skills and inferential skills, should be particularly effective at encouraging both better comprehension and metacomprehension.

4.3. Avenues for future research, methodological reflections and limitations

Together, the results from these studies provide evidence for the hypothesis that inferential processes are important for successful global absolute metacomprehension (Otero, 2002). However, more work is needed to more fully understand the role of inferential processes on metacomprehension. Ample research suggests that readers engage in inferencing during reading, aiding comprehension; but is this sufficient to yield high levels of metacomprehension? Alternatively, inferencing can occur both during reading and after the evaluation process during regulation. Clearly, this latter explanation is more likely. Still, more work is needed to shed light on the nuances of these processes and how inferential abilities influence metacomprehension in a moment-by-moment manner. Future research should explore these avenues using rigorous methods such as think-aloud protocols, especially online as students are engaged in the task itself. This approach would help clarify these nuances and provide additional empirical support for theoretical

assumptions. Likewise, future work should employ both global *and* local (i.e., item-by-item) as well as absolute *and* relative monitoring judgments, and researchers should include prospective *and* retrospective monitoring judgments to provide a clearer, finer-grained analyses of the phenomena we investigated. Furthermore, experimental studies providing training on inferential reasoning should be conducted to ascertain its effect on metacomprehension monitoring. Finally, researchers should delay performance judgments to ascertain whether the delayed judgment effect improves metacomprehension accuracy.

These studies employed a correlational design. Non-experimental designs do not permit the isolation of causal effects, and hence, the inferences and conclusions drawn from correlational data, such as regression, are not as strong. Further, some of our effect sizes were relatively small, and hence, our findings should be interpreted with the magnitude of effects in mind. Additionally, we did not collect performance metrics during reading or real-time reading processing, thereby limiting the utility of our findings during actual reading. We also acknowledge that we did not collect information on prior reading ability, and hence, this may behave as a potential confound in the study, as better readers would likely outperform poor readers by virtue of increased prior reading comprehension. In addition, the data were nested for students within classrooms, and we did not account for this nesting of the data.

Moreover, although we used objective measures such as performance and metacomprehension accuracy, it should be noted that we used a self-report survey to capture metacognitive knowledge, and hence, the survey may not have captured the full dimensionality of metacognition. Also, participants may not have been fully honest in their self-reported data. Despite these limitations, however, the sample sizes for the two studies was relatively large and provided enough statistical power to support our expectations regarding the data. Thus, we believe these studies contribute to the literature on metacomprehension monitoring and how it is influenced by factors such as type of performance (i.e., inferential and text-based).

4.4. Conclusions

The results of the present study suggest that students who have high inferential skills are also those with higher self-reported metacognitive knowledge, specifically knowledge pertaining to evaluation. These readers can efficiently repair their mental representations of a text, and exhibit higher metacomprehension accuracy in their inferential level of understanding. On the other hand, readers who operate primarily at the text-based level tend to have high metacomprehension accuracy only on a text-based level. Thus, our combined findings suggest the importance of both *inferential skills* and *metacognitive knowledge of reading strategies*, particularly as they relate to evaluation of learning. In turn, readers should be able to better understand while they read and have better regulatory abilities to guide their future efforts.

Author details

Christian Soto¹
E-mail: christiansoto@udec.cl
Antonio P. Gutiérrez de Blume²
E-mail: agutierrez@georgiasouthern.edu
ORCID ID: <http://orcid.org/0000-0001-6809-1728>
Mathew Jacovina³
E-mail: mjacovina@gmail.com
Danielle McNamara³
E-mail: dsmcnamara1@gmail.com
Nicholas Benson⁴
E-mail: nicholas_benson@baylor.edu
Bernardo Riffo¹
E-mail: bernardo@udec.cl

¹ Psycholinguistics Area, Department of Spanish, University of Concepción, Concepción, Chile.
² Department of Curriculum, Foundations, and Reading, Georgia Southern University, Statesboro, GA, USA.
³ SoLET Lab, Arizona State University, Tempe, AZ, USA.
⁴ Department of Educational Psychology, Baylor University, Waco, TX, USA.

Conflict of Interest

The authors declare that they have no conflict of interest.

Compliance with Ethical Standards

This research endeavor was approved by the Universidad Autonoma de Chile, where the data were collected.

Citation information

Cite this article as: Reading comprehension and meta-cognition: The importance of inferential skills, Christian Soto, Antonio P. Gutiérrez de Blume, Mathew Jacovina, Danielle McNamara, Nicholas Benson & Bernardo Riffo, *Cogent Education* (2019), 6: 1565067.

References

Anderson, M., & Thiede, K. (2003). Summarizing can improve metacomprehension accuracy. *Contemporary Educational Psychology*, 28, 129–160. doi:10.1016/S0361-476X(02)00011-5
Baker, L., & Beall, L. C. (2009). Metacognitive processes and reading comprehension. In S. E. Israel & G. G.

- Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 373–388). New York, NY: Routledge.
- Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology: An International Review*, 54, 199–231. doi:10.1111/apps.2005.54.issue-2
- Boekaerts, M., & Rozendaal, J. S. (2010). Using multiple calibration measures in order to capture the complex picture of what affects students' accuracy of feeling of confidence. *Learning and Instruction*, 20(4), 372–382. doi:10.1016/j.learninstruc.2009.03.002
- Bol, L., & Hacker, D. J. (2001). A comparison of the effects of practice tests and traditional review on performance and calibration. *The Journal of Experimental Education*, 69, 133–151. doi:10.1080/00220970109600653
- Bol, L., Hacker, D. J., O'Shea, P., & Allen, D. (2005). The influence of overt practice, achievement level, and explanatory style on calibration accuracy and performance. *The Journal of Experimental Education*, 73, 269–290. doi:10.3200/JEXE.73.4.269-290
- Cain, K., & Oakhill, J. V. (1999). Inference making and its relation to comprehension failure. *Reading and Writing*, 11, 489–503. doi:10.1023/A:1008084120205
- Cain, K., Oakhill, J. V., Barnes, M. A., & Bryant, P. E. (2001). Comprehension skill, inference making ability, and the relation to knowledge. *Memory and Cognition*, 29, 850–859. doi:10.3758/BF03196414
- Dunlosky, J., Rawson, K., & Hacker, D. (2002). Metacomprehension of science text: Investigating the levels-of-disruption hypothesis. In J. Otero & A. Graesser (Eds.), *The psychology of science text comprehension* (pp. 255–280). Hillsdale, NJ: LEA.
- Dunlosky, J., Griffin, T., Thiede, K., & Wiley, J. (2005). Understanding the delayed keyword effect on metacomprehension accuracy. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 31, 1267–1280.
- Dunlosky, J., & Lipko, A. R. (2007). Metacomprehension: A brief history and how to improve its accuracy. *Current Directions in Psychological Science*, 16, 228–232. doi:10.1111/j.1467-8721.2007.00509.x
- Dunlosky, J., & Rawson, K. A. (2005). Why does rereading improve metacomprehension accuracy? Evaluating the levels-of-disruption hypothesis for the rereading effect. *Discourse Processes*, 40, 37–55. doi:10.1207/s15326950dp4001_2
- Efklides, A. (2008). Metacognition: Defining its facets and levels of functioning in relation to self-regulation and co-regulation. *European Psychologist*, 13, 277–287. doi:10.1027/1016-9040.13.4.277
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive development inquiry. *American Psychologist*, 34, 906–911. doi:10.1037/0003-066X.34.10.906
- Glenberg, A. M., & Epstein, W. (1985). Calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11, 702–718.
- Glenberg, A. M., Sanocki, T., Epstein, W., & Morris, C. (1987). Enhancing calibration of comprehension. *Journal of Experimental Psychology: General*, 116, 119–136. doi:10.1037/0096-3445.116.2.119
- Griffin, T., Wiley, J., & Thiede, K. W. (2008). Individual differences, rereading, and self-explanation: Concurrent processing and cue validity as constraints on metacomprehension accuracy. *Memory & Cognition*, 36, 93–103. doi:10.3758/MC.36.1.93
- Gutierrez, A. P., & Schraw, G. (2015). Effects of strategy training and incentives on students' performance, confidence, and calibration. *Journal of Experimental Education*, 83, 386–404. doi:10.1080/00220973.2014.907230
- Hacker, D. J. (2014). Failures to detect textual problems during reading. In D. Rapp & J. Braasch (Eds.), *Processing inaccurate information* (pp. 73–92). Cambridge, MA: The MIT Press.
- Hacker, D. J., Bol, L., & Bahbahani, K. (2008). Explaining calibration accuracy in classroom contexts: The effects of incentives, reflection, and explanatory style. *Metacognition and Learning*, 3, 101–121. doi:10.1007/s11409-008-9021-5
- Hacker, D. J., Douglas, J., Plumb, C., Butterfield, E., Quathamer, D., & Heineken, E. (1994). Text revision: Detection and corrections of errors. *Journal of Educational Psychology*, 86, 65–78. doi:10.1037/0022-0663.86.1.65
- Jacobs, J. E., & Paris, S. G. (1987). Children's metacognition about reading: Issues in definition, measurement, and instruction. *Educational Psychologist*, 22, 255–278. doi:10.1080/00461520.1987.9653052
- Jiménez, V., Puente, A., Alvarado, J. M., & Arrebillaga, L. (2009). Measuring metacognitive strategies using the reading awareness scale ESCOLA. *Electronic Journal of Research in Educational Psychology*, 7(2), 779–804.
- Keener, M. C., & Hacker, D. J. (2012). Comprehension monitoring. In N. M. Norbert (Ed.), *Encyclopedia of the science of learning* (pp. 36–60). Springer.
- Kintsch, W. (1988). The use of knowledge in discourse processing: A construction-integration model. *Psychological Review*, 95, 163–182.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge: Cambridge University Press.
- Kintsch, W., & Rawson, K. A. (2007). Comprehension. In M. J. Snowling & C. Hulme (Eds.), *The science of reading* (pp. 209–226). Malden, MA: Blackwell.
- McNamara, D. S. (2004). SERT: Self-explanation reading training. *Discourse Processes*, 38, 1–30. doi:10.1207/s15326950dp3801_1
- McNamara, D. S., & Magliano, J. P. (2009). Self-explanation and metacognition: The dynamics of reading. In J. D. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 60–81). Mahwah, NJ: Erlbaum.
- McNamara, D. S. (2017). Self-Explanation and Reading Strategy Training (SERT) Improves low-knowledge students' science course performance. *Discourse Processes*. doi:10.1080/0163853X.2015.1101328
- McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, 14(1), 1–43. doi:10.1207/s1532690xci1401_1
- McNamara, D. S., Levinstein, I. B., & Boonthum, C. (2004). iSTART: Interactive strategy trainer for active reading and thinking. *Behavioral Research Methods, Instruments, & Computers*, 36, 222–233. doi:10.3758/BF03195567
- Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 26, pp. 125–173). New York, NY: Academic Press.
- Nelson, T. O., & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 1–25). Cambridge, MA: MIT Press.

- Nietfeld, J., & Schraw, G. (2002). The effect of knowledge and strategy training on monitoring accuracy. *The Journal of Educational Research*, 95, 131–142. doi:10.1080/00220670209596583
- Otero, J. (2002). Noticing and fixing difficulties while understanding science texts. In J. Otero, J. A. León, & A. Graesser (Eds.), *The psychology of the scientific text* (pp. 281–307). Mahwah, NJ: Erlbaum.
- Palincsar, A. M., & Brown, D. A. (1987). Enhancing instructional time through attention to metacognition. *Journal of Learning Disabilities*, 20, 66–75. doi:10.1177/002221948702000201
- Peronard, M. (2005). La metacognición como herramienta didáctica. *Revista Signos*, 38, 61–74. doi:10.4067/S0718-09342005000100005
- Puente, A., Jiménez, V., & Alvarado, J. M. (2009). *Escala de conciencia lectora (ESCOLA). Evaluación e intervención psicoeducativa de procesos y variables metacognitivas durante la lectura*. Madrid: Editorial EOS.
- Riffo, B., Reyes, F., Novoa, A., Véliz, M., & Castro, G. (2014). Competencia léxica, comprensión lectora y rendimiento académico en estudiantes de enseñanza media. *Literatura y lingüística*, 30, 136–165. doi:10.4067/S0716-58112014000200009
- Schraw, G. (1994). The effect of metacognitive knowledge on local and global monitoring. *Contemporary Educational Psychology*, 19, 143–154. doi:10.1006/ceps.1994.1013
- Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional Science*, 26, 113–125. doi:10.1023/A:1003044231033
- Schraw, G. (2009). Measuring metacognitive judgements. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 415–429). New York, NY: Routledge.
- Schraw, G., & Dennison, R. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, 19, 460–475. doi:10.1006/ceps.1994.1033
- Schraw, G., & Moshman, D. (1995). Metacognitive theories. *Educational Psychology Review*, 7, 351–371. doi:10.1007/BF02212307
- Schraw, G., Potenza, M. T., & Nebelsick-Gullet, L. (1993). Constraints on the calibration of performance. *Contemporary Educational Psychology*, 18, 455–463. doi:10.1006/ceps.1993.1034
- Snow, E. L., Jacobina, M. E., Jackson, G. T., & McNamara, D. S. (2016). iSTART-2: A reading comprehension and strategy instruction tutor. In D. S. McNamara & S. A. Crossley (Eds.), *Adaptive educational technologies for literacy instruction* (pp. 104–121). New York, NY: Taylor & Francis, Routledge.
- Soto, C., Rodríguez, M. F., & Gutiérrez de Blume, A. P. (2018). Exploring the meta-comprehension abilities of students with intellectual disabilities. *International Journal of Special Education*, 33(2), 233–247.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston, MA: Pearson.
- Thiede, K. W. (1999). The importance of accurate monitoring and effective self-regulation during multitrial learning. *Psychonomic Bulletin & Review*, 6, 662–667. doi:10.3758/BF03212976
- Thiede, K. W., Griffin, T., Wiley, J., & Redford, J. (2009). Metacognitive monitoring during and after reading. In J. Dunlosky, A. Graesser, & J. Hacker (Eds.), *Handbook of metacognition in education* (pp. 85–106). New York, NY: Routledge.
- Thiede, K. W., Wiley, J., & Griffin, T. D. (2011). Test expectancy affects metacomprehension accuracy. *British Journal of Educational Psychology*, 81, 264–273. doi:10.1348/135910710X510494
- Van Overschelde, J. P., & Nelson, T. O. (2006). Delayed judgments of learning cause both a decrease in absolute accuracy (calibration) and an increase in relative accuracy (resolution). *Memory & Cognition*, 34(7), 1527–1538. doi:10.3758/BF03195916
- Vieiro, P., & Gómez, I. (2004). *Psicología de la lectura*. Madrid: Pearson Educación.
- Weaver, C. (1990). Constraining factors in calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 214–222.
- Weaver, C. A., & Kelemen, W. A. (1997). Judgments of learning at delays: Shifts in response patterns or increased metamemory accuracy? *Psychological Science*, 8(4), 318–321. doi:10.1111/j.1467-9280.1997.tb00445.x
- Winne, P. H. (2004). Students' calibration of knowledge and learning processes: Implications for designing powerful software learning environments. *International Journal of Educational Research*, 41, 466–488. doi:10.1016/j.ijer.2005.08.012
- Winne, P. H., & Nesbit, J. C. (2009). Supporting self-regulated learning with cognitive tools. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 259–277). New York, NY: Routledge/Taylor & Francis Group.
- Zabucky, K. M., Moore, D., Agler, L. L., & Cummings, A. (2015). Students' metacomprehension knowledge: Components that predict comprehension performance. *Reading Psychology*, 36, 627–642. doi:10.1080/02702711.2014.950536
- Zimmerman, B. J. (2000). Attaining self-regulation: A social-cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). San Diego, CA: Academic Press.

Appendix

7. What is the composition of saliva? (*Text-based question*)

2 POINTS: Describes the three substances mentioned in the text.

1 POINT: Describe 1 or 2 substances

0: does not describe any substance.

15. What would happen if the walls of the stomach failed to contract? (*Inferential question*)

2 POINTS: Do not mix with the PH and do not degrade proteins.

1 ITEM: Just do not mix with the PH.

0: No relation to the question.



© 2019 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.

You are free to:

Share — copy and redistribute the material in any medium or format.

Adapt — remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made.

You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

No additional restrictions

You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.



Cogent Education (ISSN: 2331-186X) is published by Cogent OA, part of Taylor & Francis Group.

Publishing with Cogent OA ensures:

- Immediate, universal access to your article on publication
- High visibility and discoverability via the Cogent OA website as well as Taylor & Francis Online
- Download and citation statistics for your article
- Rapid online publication
- Input from, and dialog with, expert editors and editorial boards
- Retention of full copyright of your article
- Guaranteed legacy preservation of your article
- Discounts and waivers for authors in developing regions

Submit your manuscript to a Cogent OA journal at www.CogentOA.com

