DE ALMERÍA

Development and Validation of a Tool to Detect and Repair Text Inconsistencies

Christian Soto Fajardo[1]**, Antonio P. Gutierrez de Blume** [2]**,
Marco Antonio Contreras Castro**[1]**, Macarena Carrasco
Bernal** [1]

---

[1] Universidad de Concepción, Concepción, Chile
[2] Georgia Southern University, Georgia, USA

Chile

Correspondence: Marco Antonio Contreras Castro, Universidad de Concepción, Chile. marcontrerasc@udec.cl

Christian Soto Fajardo, Antonio P. Gutierrez de Blume, Marco Antonio Contreras Castro & Macarena Carrasco Bernal

Abstract

Introduction: Reading comprehension is an important skill in the fast pace global society in which we now live. However, reading comprehension is a rather complex process that involves the dynamic interaction of cognitive and metacognitive skills.

Method: In the present study, we investigated whether a newly-developed text inconsistency detection task adequately measured inconsistency detection and reparation. We validated the measure employing a sample of 146 undergraduate students in Chile. We also explored whether the validated measure significantly positively predicted performance on a standardized reading comprehension measure and whether it was able to adequately discriminate between proficient and poor readers.

Results: We found that the final solution of inconsistency detection task was unidimensional and explained approximately 68% of the variability in the items. Further, inconsistency detection and reparation positively significantly predicted reading comprehension performance, and it was able to successfully discriminate between proficient and poor readers. Implications for learning and educational practice are discussed.

Discussion and Conclusion: We demonstrated that our proposed text inconsistency detection test can be successfully and efficiently employed to invoke readers' skill at monitoring reading comprehension by challenging them to detect inconsistencies and pursue the metacognitive process of reparation (i.e., control).

*Keywords***:** Text inconsistency detection; text inconsistency reparation; reading comprehension; metacognition

Resumen

Introducción: La comprensión de lectura es una habilidad importante en esta sociedad globalizada y de ritmo rápido. Al mismo tiempo, es un proceso complejo que involucra la interacción dinámica de las habilidades cognitivas y metacognitivas.

Método: En el presente estudio, se investiga si una tarea de detección de inconsistencia midió adecuadamente la detección y reparación de inconsistencias en la lectura de textos breves. Se valida la medida empleando una muestra de 146 estudiantes de pregrado en Chile. También se explora si la medida validada predijo significativamente de manera positiva el rendimiento de comprensión de lectura estandarizada y si fue capaz de discriminar adecuadamente entre lectores competentes y no competentes.

Resultados: Se encuentra que la solución final de la tarea de detección de inconsistencias fue unidimensional y explicó aproximadamente el 68% de la variabilidad en los ítems. Además, la detección y reparación de inconsistencias predijeron de manera significativa el rendimiento de la comprensión de lectura, y fue capaz de discriminar con éxito entre lectores competentes y no competentes.

Discusión y conclusiones: La prueba de detección de inconsistencia puede emplearse con éxito y eficiencia para invocar la habilidad de los lectores para monitorizar el proceso de comprensión de la lectura al desafiarlos a detectar inconsistencias y perseguir el proceso metacognitivo de reparación (es decir, control).

Palabras clave: detección de inconsistencias; reparación de inconsistencias; comprensión lectora; metacognición

Christian Soto Fajardo, Antonio P. Gutierrez de Blume, Marco Antonio Contreras Castro & Macarena Carrasco Bernal

Introduction

When readers are faced with reading texts, the task of text comprehension begins, which implies a process in which a cognitive processing of language takes place. Simultaneously, the metacomprehension domain activates processes that are at a level higher than understanding, which allows readers to achieve the desired reading comprehension. The metacomprehension process ensures that understanding is being developed in an efficient way, according to coherence standards. Otero (2002) argue that these standards reflect the knowledge and beliefs of a reader about what constitutes a "good" understanding, as well as the specific objectives of the reader to read the particular text. Readers use these standards to assess their understanding and to determine if they should participate in additional comprehension processes, such as the search for additional information in episodic or semantic memory.

Reading Comprehension

When the reader does not achieve the desired reading performance it is probably because he has failed to develop a consistent mental model, an explanation given by Kintsch and van Dijk (1978) who proposed that readers represent texts on three levels: "surface code", which corresponds to the perceptual and verbal aspect of language, and includes the identification of words and the recognition of syntactic and semantic relationships between them. The second level is the "text-base", which refers to the semantic aspect of language, and it is represented by propositions. The importance of this level lies in the fact that the representation of the meaning of the sentences becomes independent of the form, as the propositional format only includes the relations between the predicates and arguments without requiring the superficial form of the text to be expressed. The third level is the "situation model", and from it we can assume that the reader constructs a representation of the specific situation posed by the text, based on his prior knowledge and the information in the text. From a psycholinguistic approach, understanding is not only about the application of linear operations performed by a passive reader, but it is "a complex and interactive process that requires the activation of a considerable amount of knowledge by the reader and of the generation of a large number of inferences" (León, 2001, p.114).

The purpose of comprehension is to construct a coherent mental representation, called a situation model (Kintsch, 1988), based on the reader's knowledge and information in the text. In other words, to achieve a coherent understanding of a text, it is necessary to generate

inferences, a task that becomes easier when we are aware of the processes that operate when we read and learn. Supporting this idea, González and Romero (2001) mention that, in general, individuals with poor reading comprehension read superficially and do not learn deeply by reading, lack knowledge or strategies necessary to identify text structures, and lack prior knowledge about the text contents. Conversely, proficient readers can generate inferences from the information delivered by the text, and hence, deeply comprehend.

Kintsch and Rawson (2005) point to the importance of the processing levels necessary to understand a text. In the first place, the work of decoding the graphic symbols that involves processes of perception, recognition of words and their roles in the propositions is necessary. Secondly, a semantic analysis of the words is necessary to give meaning to the propositions which, as a whole, will give a meaning to the text. Through the interrelation of the propositions we arrive at the microstructure of the text, built on the basis of syntactic relations, whose coherence requirement necessitates the activation of local inferences necessary to generate a coherent microstructure. Finally, it is crucial that the reader accesses the text base since it represents the semantic basic meaning of the text. However, if a reader only understands what is expressed explicitly, he achieves only a superficial understanding. To reach a deeper understanding it is necessary that the reader construct a situation model; that is, a mental model of the situation described by the text. Therefore, to achieve the understanding of the text, a reader must to use their prior knowledge to enrich the text representation. In short, to achieve a deep understanding the reader must advance strategically beyond the text and build a model of the situation referred in the text, for which the activation of inferential processes is necessary. Linked to the construction of the situation model, the reader incorporates multiple representations derived from this model among which are the spatial locations of entities and events, temporal relationships, cause and effect relations, relations between people, and direction of the process, among others.

Metacomprehension

Metacognition essentially means cognition about cognition, a term coined in the 1970s from the studies of Flavell and other contemporary authors (Flavell, 1971, Flavell, Friedrichs & Hoyt, 1970). Flavell managed to distinguish between metacognitive knowledge and metacognitive experience, in addition to explaining that metacognition refers to consciousness and control of both; that is, not only of cognitive processes but also of emotions and motivation. Even though almost fifty years have passed since researchers began studies related to

metacognition, many scientists are interested in determining and exploring the components of metacognition, and how they operate according to different cognitive processes such as attention, learning or memory. Within them we can also find the reading process, which drives a burgeoning interest due to the multiplicity of tasks that are performed when decoding a text.

With respect to the metacognitive process involved in reading, Brown (1987) distinguishes between knowledge about cognition and the regulation of cognition. Knowledge about cognition can be, "… stable … stable but fallible or late in developing." (p. 67), that thinkers have about their own cognitive processes, which generally remains relatively stable within individuals. Regulation, on the other hand, can be "… relatively unstable, and independent of age." (p. 68). The regulation of cognition refers to the activities used to regulate and supervise learning. Self-regulatory behavior can be shown in one situation, but not another, just as a child can show self-regulation behavior when an adult does not. In addition, regulation can also be affected by patterns of excitement (e.g., anxiety, fear, interest) and self-concept (e.g., self-esteem, self-efficacy); patterns that include planning before tackling a problem, monitoring activities during learning, and verifying results at the end (Brown, Bransford, Ferrara, & Campione, 1983).

The majority of metacomprehension studies focus their attention on monitoring processes because regulatory processes are of a much more complex nature to study, as they involve text production (i.e., regulation is an adjustment process that occurs during reading and, as such, is more difficult to measure reliably). The tasks for the detection of inconsistencies assume that detecting an error intentionally introduced in the text could be a way to access the evaluation made by the readers of their own understanding of the text during the construction of the meaning (Soto, Poblete & Gutiérrez de Blume, 2018b). Traditionally, monitoring and regulation are usually studied through specific tasks for each experiment through the error paradigm (Hacker, 1998). The error detection methodology allows us to address the comprehension skill through its monitoring and control actions based on self-regulation, which are deployed both in the identification of such faults and in their correction or repair.

Considering the aforementioned, when instruments are employed to measure the role of metacomprehension elements it is important to ascertain if they are calibrated and validated with respect to what is to be studied. It is relevant that this exploratory study indicates findings that may prove to be valuable for future research, as traditionally metacomprehension has been

measured through tasks and not tests that are properly validated. The main objective when a reader experiences difficulties or inconsistencies in the texts to be read is for them to generate a coherent mental representation, and therefore, force them to invoke their regulatory mechanisms to solve text inconsistencies. It is important to mention the interest that should be given to metacomprehension within the educational field, as the application of metacognitive strategies such as self-awareness and self-control enable independent learners to control their own learning and become lifelong learners (Kintsch, 1988; Kintsch & van Dijk, 1978).

Monitoring and Control of Reading

Otero and Graesser (2014) argue that the obstacles in the supervision of the comprehension of text information can be classified according to the same three levels that Kintsch (1978) and van Dijk postulate. Therefore, comprehension can be monitored at the lexical, semantic and referential levels. From this monitoring, processes of regulation of reading are triggered as an underlying operation that remains active throughout the comprehension task. In this way, the monitoring and regulation phases are implemented after the evaluation of comprehension. The regulatory process acts as the reader acts to repair or improve his understanding. This includes adjustment operations during the comprehension process, while evaluation can occur at different stages of reading, both during and/or after reading.

Regulation is a dynamic process that depends on evaluation, but it can be implemented for different reasons, such as correcting a defective understanding or deepening comprehension. Because regulatory studies have traditionally focused on error detection when reading materials that contain inconsistencies, there has been little consideration of regulation as a mechanism to understand ideas more deeply during reading (Soto, Rodríguez & Gutierrez de Blume, 2018b). Following this point of view, there must be situations in which readers decide (consciously or unconsciously) to improve their mental representation using different strategies, even when there is no inconsistency in text coherence.

Inconsistency Detection

Since the 1980s, the task of detecting errors in reading has been used. Garner (1980) conducted research to establish differences in the processing of comprehension of inconsistencies. His hypothesis suggests that there are differences between good and poor readers in the understanding of short expository texts. Inserting a semantic inconsistency in them, good readers noticed the disturbing effect of the altered text while poor readers did not.

Hacker and collaborators (Hacker, 1994) attempted to determine whether knowing about how to correct errors affects the detections. The researchers explored the following hypothesis: (a) readers know how to correct detected errors, and (b) readers who do not know how to correct errors do not detect them. If knowledge of how to correct is sufficient for detection, then we should find support for the following hypotheses: (c) readers who know how to correct errors detect them, and (d) readers do not know how to correct errors that are not detected. The researchers concluded that the knowledge of how correct errors is relevant, but it requires something additional; more specifically, the specific ability/strategie to detect errors, directing strategically his attention to specific portions of the text.

A series of error detection studies has been published, but those studies have explored the phenomenon of focusing on different types of errors (spelling, syntax, semantics, etc.) analyzing how certain variables can affect the detection capacity (working memory, attention, etc.) and utilizing different tasks and materials (For example: Larigauderie, Gaonac'h, & Lacroix, 1998). This is an issue of importance in the discipline becase, to determine regularities in error detection, we must calibrate texts and tasks well, if we want our findings to be generalizable.

Otero and Campanario (1990) investigated the effect of introducing errors or inconsistencies in a series of texts. When introducing errors or inconsistencies in the text material, some participants should presumably be able to first assess their understanding of the text as imperfect (i.e., because they recognize it as false) and, finally, solve the difficulty using different cognitive strategies. Based on the level of regulation presented, the participants were classified into three groups: 1) those who had not noticed the contradiction; 2) those who evaluated the contradiction, but did not exercise adequate regulation; and finally, 3) students who conducted an appropriate evaluation and regulation process. According to Otero and Campanario's interpretation of these different classifications of readers, evaluation is an initial process that occurs before regulation. Otero (2002), about the Inconsistency Detection, posits that the mechanism of regulation is modeled as a limitation of satisfaction in which readers evaluate the coherence of their mental representation of a text with respect to a standard. Soto and colleagues arrive at a similar idea, using comprehension tests, metacomprehension questionnaires and calculating the calibration of performance judgments of students of 7th and 8th grades (Soto, Gutierrez, Jacovina, Benson, McNamara, & Riffo, 2019a).

An investigation by Latorre and Escobar (2010) shows the relationship between aspects of the behavior of the subjects and the task of detecting errors. The findings of this study highlight that the ability to detect and correct errors can be related to the type of interaction that each reader assumes in front of the text, affecting the semantic relation that the reader generate about the text.

In working with university students, it seems relevant to work with semantic aspects of the text (on reading comprehension tasks), and to determine how students operate by integrating cognitive and metacognitive functions. Expert readers who evaluate their understanding constantly ask themselves if what they are reading makes sense. If this does not make sense, repair strategies are applied to restore understanding (i.e., metacognitive monitoring and control). In practice, the monitoring of comprehension is often reduced to detection and, if possible, to the resolution of inconsistencies such as contradictory sentences or statements in conflict with the reader's knowledge of the world (van der Schoot, Reijntjes, & van Lieshout, 2012). Therefore, comprehension monitoring measures generally involve the analysis of verbal and non-verbal responses immediately after a coherence violation (Skarakis-Doyle, 2002) or answers to questions after a text containing inconsistent or conflicting information (Zinar, 2000).

In short, the detection of inconsistencies is recognized by researchers in the field of reading comprehension as an important method, and there have been attempts to take advantage of this exploratory technique. The detection of inconsistencies allows us to determinate how the different passages of the text are processed, analyzing the semantic weight of the ideas of the sentences (Lesgold & Perfetti, 2017). This finding help determine how to intervene appropriately in both with the adequate selection and calibration of the texts, and with the interventions in the students´s reading comprehension processes (Soto, Gutierrez de Blume, Rodríguez, Asún, Figueroa, & Serrano, 2019b).

The objective of the present study is to relate comprehension skills (comprehension of texts) with those of metacomprehension (detection and repair of inconsistencies) while also validating a measurement instrument of metacomprehension skills, based on the reading of short texts in which we have intentionally incorporated information that alters its coherence. We were also interested in investigating the extent to which elements of text inconsistency detection predicted reading comprehension performance and whether they adequately discriminated between poor and proficient readers, a task necessary for diagnostic assessment.

The Present Study

Predicated on the research literature we surveyed above, we proposed the following research questions to guide this study.

Research Questions and Hypotheses

1. What is the factor structure of the newly created instrument to assess inconsistency detection in text and learners' ability to repair the inconsistencies?

Hypothesis 1: We predicted that the solution would yield one factor that would capture the process of inconsistency detection and reparation that would explain significant variability in the manifest variables.

2. Do the manifest variables of the final inconsistency detection instrument adequately predict performance on a standardized reading comprehension measure?

Hypothesis 2: We expected that the manifest variables of the final text inconsistency detection instrument would significantly and positively predict performance on a standardized reading comprehension measure.

3. Are the manifest variables of the final inconsistency detection measure able to discriminate between proficient and poor readers?

Hypothesis 3: We hypothesized that the manifest variables of the final inconsistency detection measure would adequately discriminate between poor and proficient readers.

Method

Participants

Participants recruited for this study consisted of 146 Chilean undergraduate students. 13 were male and 133 were female. Their age ranged from 19-23 years old ($M = 20.85$, $SD = 0.98$). Their majors were either Basic General Pedagogy (for those in the first year of their academic career) or Phonoaudiology (for those in the second, third, or fourth year of their academic career), and their academic standing was as follows: 55 freshman college students, 40 second-year university students, 47 third-year university students and 4 fourth-year university students. The demographic characteristics of students who participated in this study were representative of the population of undergraduate students at the university from which they were recruited.

The sample represents one drawn from convenience. Among the inclusion criteria, the participants were to be typical university students between 18 and 22 years old, who participated voluntarily. The exclusion criteria included those who had completed the test in a time frame below the average and who had failed to respond to more than 2 of the 14 items.

Instruments

Reading performance. *Lectum* is a reading comprehension test for the Chilean school system. *Lectum* is a recent reading comprehension test that successfully evaluates various cognitive components of the reading process (Riffo, Véliz, Castro, Reyes, Figueroa, Salazar and Herrera, 2011). The most relevant feature of the test is given by the fact that the instrument was developed based on a psycholinguistic model for the evaluation of reading comprehension. The main theoretical foundations that support it come from psycholinguistics, the studies of discourse and pragmatics. These foundations consider that the process of reading and its comprehension imply, among others, a reader, a text and its context. To determine the components (or sub-dimensions) of the reading process, three criteria have been elaborated, namely: 1) criteria determined by the level of processing required by the task, considered "textual" comprehension; 2) criteria determined by the context, also called "pragmatic" comporehension and 3) criteria determined by the reader and their position against the text and its context, called "critical" comprehension. From each of these criteria, sub-skills are distinguished. A relevant aspect of the model is the distinction between, on the one hand, information that is required to answer the task and that can be presented explicitly, and that which is implicit in the text, on the other. Traditionally, such a distinction has been made with the terms "literal comprehension" versus "inferential comprehension", respectively. In the present model, inferences are operations that occur practically throughout of all the process; therefore, *Lectum* is proposed to indicate, for each question, whether the information provided for its resolution is explicit or implicit in the text.

Each item correct was scored with a point, and the overall score of the test is the simple summation of the correct answers, according to norms established for each grade level. Then performance raw scores were mathematically transformed to percentiles from 0 to 100 to more readily compare it to other variables and to facilitate interpretation. Subsequently, *Lectum* gain scores were calculated by subtracting the pretest *Lectum* performance from posttest performance (i.e., posttest *Lectum* performance score – pretest *Lectum* performance score = *Lectum* gain score).

Inconsistency detection task. We developed a test of inconsistencies in which students are expected to detect both internal (i.e., linked to elements of the wording or those that are present in the text.) and external (i.e., the analysis of the memory of information inconsistent with prior knowledge) inconsistencies within sets of expository texts. This test is intended to measure learners' metacognitive monitoring and regulation of reading comprehension. First, 14 brief expository texts we developed were presented to students.[1] Selection of these 14 texts was determined through previous analyses via *Trunajod*, which collects readability indices of the texts. It is important to note that there were two texts that were employed as controls (i.e., they contained no inconsistencies), and thus, there were 12 actual texts with inconsistencies. This tool labels the texts and quantifies words, sentences, clauses, idioms and propositions. In addition to delivering the above, *Trunajod* can calculate the textual complexity in terms of clause length, sentence length, subordination index, lexical density, lexical diversity and propositional density. From all these data, we chose criteria that allowed us to calibrate the texts so that levels of difficulty were combined, and a variety of texts were available according to these characteristics. The criteria selected through the analysis of *Trunajod* were finally the following: words, sentences, notional words, propositions, sentence length, subordination index.

The text has been calibrated in terms of readability, specifically in terms of propositional density, lexical density and nominal words. This has been possible using TRUNAJOD, a technology generated on the University of Concepción, that captures most that 20 indices of textual complexity. So, we have tried to preserve a certain homogeneity in the complexity indexes, so that the texts behave in a homogeneous way in terms of lexical and syntactic complexity. Obviously, given the nature of the research, in addition to the thematic differences, the texts have been intervened in specific points to generate inconsistencies between two specific points of the text, by an inconsistency of the semantic type that affects the coherence in the mental representation of the reader.

An example:

"If we place our hands at a certain height above a hot surface, we will quickly feel an increase in temperature. The air, when heated, expands and becomes less dense than cold air. Then an upward current of hot air is produced and, in parallel to this, downward currents of cold air are generated. Thus, the cold air replaces the hot air that went down and the cycle is

renewed. This mechanism of propagation of heat is called natural convection and involves the transport of energy".

This text contains an inconsistency that makes it difficult to generate the situation model generated from reading the text (Kintsch, 1998), affecting the spatial dimension that is part of the content of the text. In the first part of the text the proposition that talks about hot air, mentions that it rises. However, in the penultimate proposition it is indicated that the hot air went down, which affects the construction of the microstructural coherence when proposing an opposite information in relation to the one that should be correlated.

Students are required to not only identify the inconsistencies but also to suggest ways in which to repair the inconsistencies to improve text coherence. We then calculated the total number of text inconsistencies correctly detected and repaired by each student as two distinct variables of interest in the present study. Each item correct was scored with 4 points overall; the detection of inconsistencies was scored with 2 points, and reparation with 2 points. The overall score of the test is the simple summation of the correct answers. Subsequently, results were mathematically transformed to a scale of 0%-100% to facilitate interpretation for research questions two and three.

Procedure

Test of inconsistencies was elaborated from 14 short texts, 12 with actual inconsistencies within them. The objective of this test is to measure the metacognitive monitoring of students and the regulation of reading comprehension. First, the 14 short texts were presented to the students. Then they were asked if there are inconsistencies within the given texts. Students were required not only to identify inconsistencies, but also to suggest ways to repair them to improve the coherence of the text. We first applied the reading comprehension measure (*Lectum*) to collect data on the inconsistency test. Next, we assigned a score to each inconsistency detected and to each inconsistency repaired. Subsequently, we proceeded to identify which text inconsistencies were detected and / or repaired.

Data Analysis

All data were screened for outliers and to ascertain if they met requisite statistical assumptions prior to data analysis. Outliers were screened using box and whisker plots and the RESIDUAL command with SPSS 23 for cases that were beyond three standard deviations from

the sample mean. In addition, we evaluated the data for normality using skewness and kurtosis statistics, homoscedasticity, multicollinearity, and homogeneity of variance, among others. No outliers were detected that would otherwise undermine the trustworthiness of the findings. Furthermore, all requisite statistical assumptions were met, and hence, data analysis proceeded without making any adjustments to the data.

Our first research objective was met by first conducting an exploratory factor analysis (EFA) using common factor extraction with an oblique rotation on the initial set of 12 texts (i.e., only those with inconsistencies and not the two control texts with no inconsistencies) which formed the manifest variables of the proposed inconsistency detection task. Standardized factor loadings, the variance the factor solution explained on the manifest variables, and our knowledge of theory were employed to more meaningfully interpret the final factor structure. To more deeply explore the latent dimensionality of the instrument, we conducted post hoc analyses to justify the final set of manifest variables as well as statistical analyses related to certain parameters of the task (e.g., clause length and other text characteristics).

The second research question was answered by conducting an ordinary least squares regression, with the final manifest variables of the inconsistency detection task serving as predictors and reading performance on a standardized reading performance test (i.e., *Lectum*) serving as the criterion. The final research question was addressed by conducting a discriminant function analysis (DFA). The manifest variables of the final inconsistency detection task served as the independent variables and reading performance skill (poor, proficient) served as the categorical dependent variable. The division of poor versus proficient readers was accomplished via the median split procedure such that those whose percentile was at or below 69 was coded as a "poor reading comprehender" and those at or above 70 were coded as "proficient reading comprehenders". Canonical correlations standardized canonical discriminant function coefficients, Wilk's λ and its inverse to determine explained variance of the solution, classification results, and leave-one-out classification results were interpreted for this purpose, along with the χ^2 test and its associated statistical significance.

The effect size reported for the regression analysis was the squared multiple correlation coefficient (R^2) and its adjusted value ($R^2_{adjusted}$) and for the DFA it was η^2. Cohen (1988) provided the following interpretive guidelines for these effect sizes: R^2—.010-.2499 as small; .250-

.499 as medium; and $\geq$ .500 as large; η^2—.010-.059 as small; .060-.1399 as medium; and $\geq$ .140 as large.

Results

Factor Analysis and Post Hoc Analyses

Upon completion of test development, we piloted the test as part of a separate study (Soto, Gutierrez de Blume, Rodriguez, Asun, Figueroa, & Serrano, under review). Slight adjustments were made to the wording of the various texts as part of this piloting process. Next, we administered the revised texts to the present sample.

Exploratory Factor Analysis (EFA)

We first conducted an EFA using common factor extraction (principal axis factoring) with a promax rotation to examine the latent factor structure of the inconsistency detection test and to remove any potentially problematic manifest variables (i.e., texts). The Kaiser criterion (Kaiser, 1960), scree plots, factor loadings, and total variance accounted by the factor solution were used to determine the final factor structure. Items with loadings higher than .30 were classified as significant (Hair, Black, Babin, Anderson, &Tatham 2006; Tabachnick & Fidell, 2013).

The initial solution yielded three factors, two with only one manifest variable loading unto each respectively. Furthermore, several of the manifest variables failed to load on any factor (i.e., texts 2, 5, 6, and 12). This initial solution only accounted for 43.96% of the total variance, with weak inter-factor correlations (rs $\leq$ .29). Thus, we proceeded to eliminate items 2, 5, 6, and 12 from the measure.

The solution with the remaining eight items was improved, yielding two factors and accounting for 56.32% of the variance in the solution. However, the second factor in this solution had one item, and four additional items did not load on any factor (7, 8, 10, and 11). As in the previous solution, the inter-factor correlation was modest (r = .21). Based on the lack of interpretability of the second factor, we opted to further eliminate items 7, 8, 10, and 11.

The third solution with the remaining four items (1, 3, 4, and 9) yielded a single factor accounting for 69.87% of the variance in the solution. Factor loadings of the four items ranged

from .499 to .921. Given the interpretability of this one-factor solution, we retained this as the final factor structure. Below we elaborate more specifically why these four manifest variables were superior, and hence, retained in the final solution.

Justification of the Final One-Factor Solution

Clause length. We proceeded to analyze the characteristics of the four texts as well as specific aspects of the inconsistencies that allowed us to determine the distinctive characteristics of these items. TRUNAJOD was used to analyze the readability indexes of the 12 original items of the test (i.e., those with inconsistencies). Next, we compared the behavior of the final texts—namely, clause length—that were part of the final instrument (1, 3, 4, 9) and those that were not (2, 5, 6, 7, 8, 10, 11, and 12). From this statistical analysis it was possible to determine that in the clause length index, the two groups of texts show significant differences, $t(11) = 2.93$, $p = .01$, *Cohen's d* = -1.44, with the four texts that were included in the final instrument ($M = 7.90$, $SD = 2.24$) demonstrating shorter clause length than those not included ($M = 13.76$, $SD = 5.26$).

These results suggest that the difference between clause length is quite sizable. The clause length is a syntactic index that has long been a relevant variable in the complexity of texts. In fact, several decades ago, Hunt (1970) argued that the ability to combine sentences forming structures of increasing complexity is considered a critical variable in syntactic maturity. In this sense, a text will be more complex in its processing when the sentences that comprise it are formed, in turn, by a greater number of constituent elements—these elements being not only words, but clauses in relations of subordination. The clause length index alludes to the number of words contained in a clause within itself, which implies that, within these units, there is a variety of elements that affect the processing of sentences in which such a clause is embedded. Such an effect, according to Campos, Contreras, Riffo, Véliz., & Reyes (2014), has important consequences for reading performance. The simple clauses, with few elements within their complements or arguments, would demand scarce resources for their cognitive processing while the more extensive clauses or with greater elements would demand greater cognitive resources. One of the crucial resources that is relevant in the processing of clauses and sentences is working memory. Obviously, if the clauses contain more elements, working memory is overloaded, leaving unprocessed elements during reading.

If we assume that the items that were not part of the final instrument (i.e., 2, 5, 6, 7, 8, 10, 11, and 12) presented greater syntactic complexity, we can infer that, when processing a complex text, performance will deteriorate. In addition, understanding is only one of the tasks that operate in this inconsistency test; the other is, obviously, the detection of inconsistencies in the text and their eventual repair.

Given such a set of tasks, the fact that the text presents greater complexity turns out to be a challenge for readers, especially when the tasks require resources both from the cognitive domain (i.e., to understand), and from the metacognitive domain (i.e., detect and repair). Given this scenario, the length of the clause is especially important because only clauses of an extension of approximately eight average words—and not greater than eight—would be adequately related to performance in reading comprehension. According to McNamara and her colleagues (McNamara, Graesser, McCarthy, & Cai, 2014), texts with fewer clauses, fewer words per sentence, and fewer words before the main verb are easier to read, and thus, will help increase syntactic simplicity, a critical index in text comprehension. Because of this analysis, we determined that both the clause length and the nature of the inconsistency in the text played a central role in these items.

Characteristic of the inconsistency in the text. In addition, we conducted an in-depth review of the texts used in the test of inconsistencies in *semantic-discursive* terms. Findings demonstrated that those texts that formed part of the final instrument (i.e., 1, 3, 4, 9) had the following pattern of regularity: the reference was part of the main idea of the text, and hence, the information that was inconsistent in the text was relevant to it, as it directly referred to the subject/object of the text itself.

Conversely, when the inconsistency was present in a secondary idea of the text, the inconsistency was not detected so easily. In this regard, Tapiero (2007) mentions that understanding a text requires representing the situation it describes. Tapiero argues that if people cannot imagine a situation in which some people are able to evoke certain relevant experiences of what they read, then understanding a text would not be easy.

Moreover, according to Tapiero (2007), understanding the textual relationships between the event or events described by the text, locally and globally, allows us a thorough understanding that aids in the development of an accurate mental representation of interrelated ideas within

texts. If it is not possible to represent a secondary aspect of a text, it is plausible to understand it at least in a general way. However, if the general idea of this is not understood, it will be very difficult to obtain a relevant mental representation of the text. These assumptions lead us to infer that it would be much easier to detect inconsistencies when said abnormality of the text affects the general idea or the centrally represented referent because this incongruence generates disruption with mental coherence. On the other hand, if the inconsistency occurs over a secondary feature, it could be omitted, eliminated from cognition, or the learner could use another mechanism to decrease the weight of that piece of information in the generated mental representation (Otero & Kintsch, 1992; van Oostendorp, 2002).

Descriptive statistics for the initial EFA results are presented in Table 1; Table 2 contains the communalities (initial and after extraction) and the factor loadings of the final one-factor solution. Descriptive statistics of the composite scores by group—poor and proficient readers—and internal consistency reliability coefficients, Cronbach's α, can be found in Table 3. Zero-order correlations for the variables of interest are presented in Table 4.

Table 1. *Descriptive Statistics of the 12 Initial Manifest Variables of the Inconsistency Detection Measure*

Manifest Variable	*M*	*SD*	Communality	
			Initial	Extraction
Item 1	0.58	0.91	0.55	0.73
Item 2	0.57	0.90	0.30	0.22
Item 3	0.64	0.94	0.59	0.63
Item 4	0.84	0.99	0.50	0.56
Item 5	1.29	1.54	0.34	0.21
Item 6	0.12	0.46	0.21	0.23
Item 7	0.85	0.99	0.25	0.27
Item 8	0.01	0.08	0.12	0.07
Item 9	0.66	0.94	0.60	0.83
Item 10	0.70	1.21	0.33	0.25
Item 11	0.64	0.94	0.30	0.22
Item 12	0.33	0.74	0.29	0.27

N = 146

Table 2. *Communalities and Factor Loadings of the Final Four Manifest Variables of the Inconsistency Detection Measure*

Manifest Variables	Communalities		Factor Loadings
	Initial	Extraction	
Item 1	0.47	0.56	0.74
Item 3	0.55	0.85	0.92
Item 4	0.49	0.54	0.51
Item 9	0.55	0.60	0.64

N = 146

Table 3. *Descriptive Statistics and Internal Consistency Reliability Coefficients of the Text Inconsistency Detection and Repairing of Text Inconsistencies Scales Between Proficient and Poor Readers*

Subscale	Proficient Readers (*n* = 71)		Poor Readers (*n* = 75)		α
	M	*SD*	*M*	*SD*	
Text Inconsistency Detection	47.89	37.49	21.00	29.07	0.79
Text Inconsistency Repairing	32.75	33.43	14.00	25.75	0.77

N = 146

Table 4. *Zero-Order Correlation Coefficients (Pearson's r) of Reading Comprehension Performance and Text Inconsistency Detection and Repairing of Text Inconsistencies Scales*

Variable	1	2	3
1. Lectum	-	.45*	.42*
2. Text Inconsistency Detection		-	.79*
3. Text Inconsistency Repairing			-

N = 146   * *p* < .01

Prediction of Lectum Performance

We conducted a standard/simultaneous ordinary least squares regression to answer the second research question, with text inconsistency detection and text inconsistency repairing serving as predictors and *Lectum* performance serving as the criterion. Results indicated that the model was statistically significant, $F_{(2,143)} = 17.49$, $p < .001$, $R^2 = 324$ ($R^2_{adjusted} = .308$). Both text inconsistency detection ($b = 0.396$ [CI$_{95\%}$ = 0.204, 0.588]; $\beta = 0.558$) and text inconsistency repairing ($b = 0.219$ [CI$_{95\%}$ = 0.103, 0.342]; $\beta = 0.284$) were significant positive predictors of reading comprehension performance.

Discriminant Function Analysis

A direct discriminant function analysis (DFA) was conducted to examine the loadings of participants' performance on inconsistency detection and their skill at repairing said inconsistencies between poor and proficient readers. Classification variables were the composite score of the inconsistency detection test (mean of items 1, 3, 4, and 9) and the composite score on participants' skill at repairing text inconsistencies. The pooled within-groups correlation for text inconsistency detection was $r = .78$.

Because two groups were used for the classification, the analysis calculated only one discriminant function (employing the total number of discrete groups − 1formula: 2 groups − 1= 1 discriminant function), with a $\chi^2 (2) = 21.77$, $p < .001$. The discriminant function accounted for 35.7% of between-group variability (1 - Wilk's $\lambda = 1 - .643 = .357$).

The discriminant function optimally discriminated between poor and proficient readers. Only standardized canonical discriminant function coefficients greater than .50 were interpreted as meaningful. The loadings of predictors into discriminant functions reported in Table 6 (i.e., standardized canonical discriminant function coefficients and structure matrix coefficients) suggested that the best predictor for distinguishing between poor and proficient readers was performance in correctly detecting inconsistencies in text. Even though accurately repairing inconsistencies was the inferior predictor for discriminating between the two groups, it still exceeded the .50 threshold, and thus, is also a meaningful discriminator between poor and proficient readers. As reported in Table 3, proficient readers significantly outperformed poor readers on both inconsistency detection and skill in correctly repairing said inconsistencies.

Finally, the classification results of the total sample of 146 participants was calculated. The number of original cases correctly classified were the following: 100 of the full sample, 51 of the poor readers, and 49 of the proficient readers, which is approximately 68.5% accuracy. The cross-validation of the classification results was done with the leave-one-out (jackknife) classification. This cross-validated classification yielded 72.2% of correctly classified cases (105 for the entire sample; 54 for the poor readers group; and 51 for the proficient readers group). Both the original and cross-validated classification results were better than the 50% by chance alone. However, z-tests were conducted to examine whether the original and cross-validated classification results were significantly better than by chance alone. Both the original, $z = 3.82$, $p < .001$, and the cross-validated, $z = 4.47$, $p < .001$, classification were statistically significantly better than by chance alone.

Summary of Findings

Results of the EFA showed that the one factor solution, explaining a sizable amount of the variance in the manifest variables (~68%) was the best solution. Additional analyses revealed that the four-item solution was most appropriate based on several key indexes of texts such as clause length and characteristics of the inconsistencies themselves. Perhaps most importantly, the final factor structure of the inconsistency detection measure was consistent with theoretical assumptions of text comprehension (Kintsch, 1998; McNamara & Magliano, 2009). In addition, inconsistency detection and reparation were both significant positive predictors of reading comprehension performance, accounting for approximately 30% of its variability. Finally, both inconsistency detection and reparation significantly discriminated between poor and proficient readers. Taken together, these findings highlight the utility and theoretical cohesion of the final inconsistency detection measure as well as its ability to significantly predict reading comprehension performance.

Discussion and conclusion

Implications for Theory, Measurement and Practice

Reading comprehension is an important skill for learners across the developmental trajectory. Mastery of fundamental skills necessary for text comprehension becomes even more pressing for adult learners who are expected to be proficient at reading to be successful contributors to society. Based on our understanding of the theoretical assumptions and the observed empirical data (i.e., high factor loadings and explained variance of the final one-factor solution),

we believe that the test of inconsistency detection with the four texts will be instrumental in assisting educators to train learners to accurately detect inconsistencies in text and invoke metacognitive monitoring and regulatory skills to correctly repair the inconsistency. Being able to detect and repair text inconsistencies is paramount to building a precise mental representation of the information in the text, and hence, proficient reading comprehension.

Our research findings demonstrated that our inconsistency detection measure appropriately discriminates between proficient and poor readers, an essential component for any evaluation and diagnostic measure. Research on differences between poor and proficient readers is not new. Extensive extant research has compared the reading performance between poor and proficient readers at various stages of the developmental trajectory, including children (e.g., Cain et al., 2004; Nation & Snowling, 1998), adolescents (e.g., Soto et al, 2018b; Soto et al., 2018b), and adults (e.g., McNamara, 2004; Soto et al., 2018a), and concluded that poor readers stand to benefit from a variety of cognitive strategies and metacognitive skills such as accurate metacognitive monitoring and bridging strategies that aid in improving text cohesion (e.g., McNamara & Magliano, 2009; McNamara, 2004; Soto et al., 2018a; Soto et al., 2018b). Thus, our proposed inconsistency detection measure can assist researchers and practitioners (e.g., classroom teachers) to find gaps in reading comprehension, particularly in the service of poor readers. Tailored strategy training interventions can then be developed to improve the reading comprehension of these struggling readers.

The inconsistency detection and repair task allows students to elicit actions such as re-reading and correcting after focusing on the error. With the construction and validation of the inconsistency detection test, we can surmise that there is a sequence of metacognitive monitoring and control strategies that would include actions such as detecting, identifying and correcting, all essential tasks necessary for deep reading comprehension, especially with complex texts. This work has allowed us to initiate a discussion on the understanding of the process of self-regulation and the implications that present elements such as the detection of semantic inconsistencies, their correction, and the consequent actions that permit accurately understanding a text.

Students were subsequently asked whether inconsistencies exist within the given texts. The short texts presented in the inconsistency test have the communicative intention of objectively reporting on their topics, and their grammatical realization is based on the concatenation

of declarative sentences. Most of the texts are scientific or technical and, therefore, constitute informative microtexts that present some level of complexity. The texts are about the process of temperature, the properties of the sound, the organisms and environment, the planets, cells, cytogenetics, hormones, some physical properties, location about countries, etc. As the texts of science, the majority assumes an expository rhetorical structure presenting objects or procedures related to some scientific aspect are explained or described (Graesser & Otero, 2002).

Avenues for Future Research

Our inconsistency detection measure evaluates an individual's skill not only to comprehend texts but also to detect inconsistencies within said texts and how to repair them. However, it does not assess learners' metacognitive skills. Hence, research is needed to examine the relations between inconsistency detection and reparation and other metacognitive skills such as planning, monitoring, and evaluation. Researchers can, for instance, use the inconsistency detection task with absolute and relative metacognitive monitoring judgments. Additionally, work is needed to examine potential interventions aimed at improving readers' skill to successfully detect text inconsistencies and repair them. Such research could help inform classroom practice in kindergarten through third grade, in which reading comprehension is the focus. Moreover, our sample included only young adult learners. If we are to truly and more deeply understand the development of reading comprehension performance across the lifespan, more research incorporating children, adolescents, and middle-aged and older adults is needed to provide evidence on the stability of our findings for various developmental levels. Finally, more research is needed to explore how measures researchers develop are invariant across cultures. It is not often that researchers investigate the influence of culture on the various constructs under investigation. Thus, research is needed to examine the invariance of our proposed text inconsistency detection task.

On the other hand, the sequence of self-regulatory actions and their implementation are specific aspects that must be studied to determine the existence of combinations of metacomprehension actions that lead to successful performance in text comprehension. This implies undertaking methodological challenges that allow us to account for other aspects of metacomprehension to more deeply examine the metacognitive processes related to the online composition of reading.

Methodological Reflections and Limitations

Even though we attempted to develop and validate a comprehensive tool to help readers detect inconsistencies in text and to repair those inconsistencies, our study does have limitations worth noting. Methodologically, we recruited a convenience sample of young adult learners. Thus, our research findings may not be readily generalizable to other samples of this population of learners. Also, the inferences and conclusions drawn from our data need to be interpreted with caution due to the non-experimental nature of our research design. In addition, we did not include additional measures of reading comprehension achievement, which would have helped triangulate our results. Finally, even though our study included objective measures, it would have been useful to include subjective metacognitive measures such as the Metacognitive Awareness Inventory (MAI; Schraw & Dennison, 1994) or the Metacomprehension Inventory (MI; Soto et al., 2018a) to relate how perceptions are congruent or incongruent with objective measures. Despite these limitations, however, our study contributes substantially to the understanding of the dynamic relation between reading proficiency and metacognition, and the instrument allows generalizations regarding the metacomprehension behavior of an age segment as well as to generate metacomprehension profiles from the results.

Conclusion

Understanding reading comprehension more deeply involves employing different types of measurement and different levels of analysis (McNamara, 2011). Therefore, we developed, piloted, and validated a test of text inconsistency detection and reparation using different science microtexts. Considering the complexity of the relation between metacognition and reading, we uncovered some interesting results. We found that our proposed text inconsistency detection measure is unidimensional and explains a significant amount of variance in the manifest variables. Moreover, we discovered that detection of text inconsistencies and their reparation positively and significantly predicted reading comprehension performance and that detection and reparation significantly discriminated between poor and proficient readers. This fact is of special importance because, from the Construction-Integration model (Kintsch, 1998), the level of reading comprehension assumes some specific mental representation according to the different reading levels (McNamara & Magliano, 2009). Presumably, there is a reciprocal and interactive influence from the cognitive to metacognitive and from metacognitive to cognitive levels in this process to aid in successfully configuring and consolidating a stable representation of the text.

Of special significance, we demonstrated that our proposed text inconsistency detection test can be successfully and efficiently employed to invoke readers' skill at monitoring reading comprehension by challenging them to detect inconsistencies and pursue the metacognitive process of reparation (i.e., control). Thus, classroom teachers can use our measure as a diagnostic and evaluative tool regarding metacognitive processes such as reading comprehension, monitoring, and control.

As an example, COMPRENDE was developed as an educational technology that teaches and trains reading comprehension strategies, proposing cognitive and metacognitive task, with student of 5th to 7th grade. The cognitive strategies taught are bridging inference, global comprehension, contextual vocabulary and idea-integration. On the metacognitive dimension the artificial intelligence generated feedback about the performance on the tasks, giving retrospective and prospective feedback. Additionally, a series of tasks are available where the students were tasked to identify a series of internal and external inconsistencies regarding different microtexts and generate the respective reparation on the specific incoherent piece of inconsistent information. After the response, the system provides feedback about the quality of the answers, showing where the error on the texts lies. COMPRENDE was shown to significantly improve reading comprehension performance for poor readers.

The inconsistency detection process is important to help enhance the accuracy of metacognitive monitoring, aiding in the regulation of reading comprehension. Thus, it is important to evaluate the level of the dimensions by a validated method.

Funding

References

Brown, A. (1987). Metacognition, executive control, self-regulation, and other more mysterious mechanisms. In Weinert F.E & Klue, R.H (Eds.) *Metacognition, motivation, and understanding* (pp. 65-116). Hillsdale, NJ: Erlbaum.

Brown, A. L., Bransford, J. D., Ferrara, R., & Campione, J. (1983). Learning, remembering and understanding. In J. H. Flavell, & E. M. Markman (Eds.), *Handbook of child psychology: Vol. 3. Cognitive development* (pp. 77-166). New York: Wiley.

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology, 96*, 31-42.

Campos Saavedra, D., Contreras Carmona, P., Riffo Ocares, B., Véliz, M., & Reyes Reyes, A. (2014). Complejidad textual, lecturabilidad y rendimiento lector en una prueba de comprensión en escolares adolescentes. *Universitas Psychologica*, *13*(3).

Flavell, J. H. (1971). First discussant's comments: What is memory development the development of? *Human development, 14(4)*, 272-278.

Flavell, J. H., Friedrichs, A. G., & Hoyt, J. D. (1970). Developmental changes in memorization processes. *Cognitive Psychology, 1(4)*, 324-340.

Garner, R. (1980). Monitoring of understanding: An investigation of good and poor readers awareness of induced miscomprehension of text. *Journal of Reading Behavior*, *12*, 55-63.

González, M.J. & Romero, J.F. (2001). Intervención psicoeducativa en comprensión lectora. En Romero, F. & González, M.J. (Eds). *Prácticas de comprensión lectora. Estrategias para el aprendizaje* (pp.17-38). Madrid: Alianza Editorial.

Graesser, A. C., & Otero, J. (2002). The psychology of science text comprehension. In Otero, J.; León, J. A.; Graesser, A. C. (eds.). *The psychology of science and text comprehension* (pp. 1-18). Mahwah, NJ, Lawrence Erlbaum.

Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E. and Tatham, R.L. (2006) *Multivariate Data Analysis. Vol. 6*, Pearson Prentice Hall, Upper Saddle River.

Hacker, D. J. (1998). Definitions and empirical foundations. In Hacker, D. J., Dunlosky, J., & Graesser, A. C. (Eds.). *Metacognition in educational theory and practice. Mahwah* (pp. 15-38), NJ: Erlbaum.

Hacker, D. J., Plumb, C., Butterreld, E. C., Quathamer, D. & Heineken, E. (1994). Text revision: Detection and correction of errors. *Journal of Educational Psychology*, *86*, 65-78.

Hunt, K. W. (1970). How little sentences grow into big ones. In M. Lester (Ed.), *Readings in applied transformational grammar* (pp. 193-201). New York: Holt, Rinehart and Winston.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and psychological measurement*, *20(1)*, 141-151.

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

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95(2)*, 163-182.

Kintsch, W. & Rawson, K.A. (2005). Comprehension. In M.J. Snowling, & C. Hulme, (Eds.), *The science of reading: A handbook* (pp. 209–226). Malden MA: Blackwell.

Kintsch, W., & van Dijk, T. (1978). Toward a Model of Text Comprehension and Production. *Psychological Review, 85(5)*, 363 – 394.

Larigauderie, P., Gaonac'h, D. Lacroix, N. (1998). Working memory and error detection in texts: What are the roles of the central executive and the phonological loop? *Applied Cognitive Psychology, 12*, 505-527.

Latorre Velásquez, D.C. & Escobar-Melo, H. (2010). Tarea Cognitiva de Detección y Corrección de Errores en la comprensión de Textos Narrativos en Niños de Básica Primaria. *Universitas Psychologica, 9 (3)*, 863-878

León, J. A. (2001). Las inferencias en la comprensión e interpretación del discurso: un análisis para su estudio e investigación. *Revista signos, 34*(49-50), 113-125.

McNamara, D.S. (2004). SERT: Self-explanation reading training. *Discourse Processes*, *38*, 1-30.

McNamara, D. S. (2011). Measuring deep, reflective comprehension and learning strategies: challenges and successes. *Metacognition and Learning*, *6*(2), 195-203.

McNamara, D. S., Graesser, A. C., McCarthy, P., & Cai, Z. (2014). *Automated Evaluation of Text and Discourse with Coh-Metrix*. Cambridge, UK: Cambridge University Press.

McNamara, D. S., & Magliano, J. (2009). Chapter 9 Toward a Comprehensive Model of Comprehension. Psychology of Learning and Motivation. *Advances in Research and Theory, 51*, 297-384.

Nation, K., & Snowling, M. J. (1998). Semantic processing and the development of word-recognition skills: Evidence from children with reading comprehension difficulties. *Journal of memory and language*, *39(1)*, 85-101.

Otero, J. (2002). Noticing and fixing difficulties while understanding science texts. In Otero, J.; León, J. A.; Graesser, A. C. (eds.). (2002). *The psychology of science and text comprehension* (pp. 1-18). Mahwah, NJ, Lawrence Erlbaum.

Otero, J. C., & Campanario, J. M. (1990). Comprehension evaluation and regulation in learning. from science texts. *Journal of Research in Science Teaching, 27(5)*, 447-460.

Otero, J., & Kintsch, W. (1992). Failures to detect contradictions in a text: What readers believe versus what they read. *Psychological Science*, *3(4)*, 229-236.

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19(4)*, 460-475.

Skarakis-Doyle, E. (2002). Young children's detection of violations in familiar stories and emerging comprehension monitoring. *Discourse Processes, 33*, 175–197.

Riffo, B., Véliz, M., Castro, G., Reyes, F., Figueroa, B., Salazar, O., & Herrera, M. O. (2011). *LECTUM. Prueba de comprensión lectora*.

Soto, C., Gutierrez de Blume, A., Jacovina, M., McNamara, D., Benson, N., & Riffo, B. (2019a). Reading comprehension and metacognition: The importance of inferential skills. *Cogent Education*, *6*(1), 1565067.

Soto, C., Gutierrez de Blume, A., Rodríguez, M. F., Asún, R., Figueroa, M., & Serrano, M. (2019b). Impact of Bridging Strategy and Feeling of Knowing Judgments on Reading Comprehension Using COMPRENDE: an Educational Technology. *TechTrends*, 1-13.

Soto, C., Gutierrez de Blume, A. P., Asún, R., Jacovina, M., & Vásquez, C. (2018a). A Deeper Understanding of Metacomprehension in Reading: Development of a New Multidimensional Tool. *Frontline Learning Research 6(1)*, 31-52.

Soto, C., Rodríguez, M. F., & Gutierrez de Blume, A. P. (2018b). 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*. Pearson, Boston.

Tapiero, I. (2007). *Situation models and levels of coherence: Toward a definition of comprehension*. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

van der Schoot, M., Reijntjes, A., & van Lieshout, E. C. (2012). How do children deal with inconsistencies in text? An eye fixation and self-paced reading study in good and poor reading comprehenders. *Reading and Writing, 25(7)*, 1665-1690.

van Oostendorp, H. (2002). Updating mental representations during reading science text. In J. Otero, J. A. Leon, & A. C. Graesser (Eds.), *The psychology of science text comprehension* (pp. 309–329). Mahwah, NJ: Erlbaum.

Zinar, S. (2000). The relative contributions of word identification skill and comprehension-monitoring behavior to reading comprehension ability. *Contemporary Educational Psychology, 25*, 363–377.