Sanjosé, Vicente; Gómez-Ferragud, Carlos B.; Verdugo-Perona, José J.; Solaz-Portolés, Joan J.
Testing a Model for the Monitoring of Worked-out Algebra-Problem Examples: From Behaviours to Outcomes on a Math Task
Psicología Educativa, vol. 28, no. 2, 2022, pp. 141-149
Colegio Oficial de la Psicología de Madrid

DOI: https://doi.org/10.5093/psed2021a25

Available in: https://www.redalyc.org/articulo.oa?id=613772153007

Psicología Educativa (2022) 28(2) 141-149

Psicología Educativa

https://journals.copmadrid.org/psed

Testing a Model for the Monitoring of Worked-out Algebra-Problem Examples: From Behaviours to Outcomes on a Math Task

Vicente Sanjosé, Carlos B. Gómez-Ferragud, José J. Verdugo-Perona, and Joan J. Solaz-Portolés

University of Valencia, Spain

ARTICLE INFO

Received 6 May 2020
Accepted 14 June 2021
Available online 22 September 2021

Keywords:
Psychological models
Math learning
Monitoring worked-out examples
Behaviour-outcome association
Online data

A B S T R A C T

This study aimed at testing an extension of a theoretical model for the metacognitive monitoring mechanism implied in the detection of inconsistencies when the information provided includes abstract symbols in addition to plain text. Ninety-four postgraduates of STEM specialities were asked to read a worked-out algebra-problem example and to report any incoherence, inconsistency, or error detected in the statement or in the solving procedure. A set of model-inspired indexes was defined to describe participants' behaviour along the task. The Read & Answer software was used to record online individual processing data and participants' reports. Results supported model predictions. Indexes correctly predicted participants' outcomes in the task with high accuracy. Specific students' behaviours could be associated to observed task outcomes with sufficient reliability within the limitations of the study. In addition, algebra processing was compared with plain text processing.

La validación de un modelo para el control de la comprensión de problemas de álgebra resueltos: de los comportamientos a los resultados en una tarea de matemáticas

Palabras clave:
Modelos psicológicos
Aprendizaje de matemáticas
Control de la comprensión de problemas resueltos
Asociación comportamiento-resultado
Datos online

R E S U M E N

Este estudio tiene como objetivo probar una extensión de un modelo teórico para el mecanismo de control metacognitivo que sirve para la detección de inconsistencias cuando la información proporcionada incluye símbolos abstractos además de texto natural. Se pidió a 94 postgraduados de especialidades STEM que leyeran un ejemplo resuelto de un problema de álgebra y que informaran sobre cualquier incoherencia, inconsistencia o error detectado en el enunciado o en el procedimiento de resolución. A partir de un modelo teórico se definió un conjunto de índices para describir el comportamiento de los participantes a lo largo de la tarea. Se utilizó el software Read & Answer para registrar online los datos de procesamiento individual y los informes de los participantes. Los resultados confirman las predicciones del modelo. Los índices predicen correctamente los resultados de los participantes en la tarea con gran precisión. Los comportamientos específicos de los alumnos podrían asociarse a los resultados observados de la tarea con suficiente confiabilidad dentro de las limitaciones del estudio. Además se ha comparado el procesamiento del álgebra con el procesamiento del texto natural.

Mechanisms explaining how the human mind operates while doing complex tasks are needed to consistently improve teaching and learning so as to avoid trial-and-error attempts. Teachers could have the possibility of more effective interventions if they knew what piece of a mental mechanism is causally associated to certain students' difficulties. Students' behavioural data along a task have been obtained in prior studies using modern devices (Cerdán et al., 2011; Mañá et al., 2017; Salmerón et al., 2017; Sánchez & García-Rodicio, 2013). Behavioural data can thus be causally related to the observed task outcomes. The present study is in line with these studies.

Among the different factors statistically associated to school learning (Wang et al., 1993), metacognitive knowledge and skills have proven to be strong predictors of academic success (Hartman, 2001; Koçak & Boyaci, 2010; Young & Fry, 2008). The relationship between metacognitive knowledge and skills, and deep approach of learning has been highlighted by García et al. (2015). There is an agreement on the benefits of improving students' metacognitive skills by explicit instruction (Desoete & De Craene, 2019), but the mental mechanisms explaining the success of these instructional approaches are poorly known.

Cite this article as: Sanjosé, V., Gómez-Ferragud, C. B., Verdugo-Perona, J. J., & Solaz-Portolés, J. J. (2022). Testing a model for the monitoring of worked-out algebra-problem examples: From behaviours to outcomes on a math task. *Psicología Educativa, 28*(2), 141-149. https://doi.org/10.5093/psed2021a25

Correspondence: vicente.sanjose@uv.es (V. Sanjosé).

Interesting advances arise from the field of artificial intelligence. Different architectures for models of metacognition incorporate human characteristics, as does, for instance, the CLARION cognitive architecture (Sun et al., 2006) or the MIDCA dual-cycle model (Cox et al., 2016). However, human mind processing is not based on strict "If-Then" rules (Otero et al., 2008), the basic underlying mechanism in most of these architectures, but on flexible distributed processing which is being revealed in recent research (Qiu et al., 2018; Yeung & Summerfield, 2014). Among other findings, neural deep research has shown that metacognitive skills are, in part, domain specific and their effectiveness depend on the task being performed (Morales et al., 2018). This suggests using tasks associated with specific subjects to study metacognitive skills.

In mathematical education, some studies devoted to increase students' metacognitive competence have been developed (Baten et al., 2017; Desoete & De Craene, 2019), using a variety of math tasks as addition and subtraction (Throndsen, 2011), proportional reasoning (Modestou & Gagatsis, 2010), or problem solving (Cornoldi et al., 2015; Lucangel et al., 2019; Özsoy & Ataman, 2017). However, most studies conducted in the classroom lack a clear theoretical basis.

Some models of psychological nature have been also proposed to explain students' metacognitive behaviour in particular school tasks, as the PREG model by Otero and Graesser (2001) or the obstacle-goal model by Otero (2009) (see also Sanjosé & Torres, 2014). Otero and Kintsch (1992) proposed a mechanism to account for monitoring comprehension of textual information based on Kintsch's (1998) Construction-Integration model, one of the best-known models for comprehension.

Aims and Overview

The present study aimed at testing an extension of a theoretical model for the metacognitive monitoring mechanism, the one proposed by Otero and Kintsch (1992), when the information to be monitored is a worked-out algebra-problem example. This kind of learning material, as others in math education, includes abstract symbols in addition to plain text. To achieve this objective, different actions were performed:

- A set of behavioural indexes were defined, inspired by Otero and Kintsch's (1992) mechanism (the model onwards), to characterize the process of monitoring worked-out algebra-problem examples for understanding.

- Experimental data were obtained to test the model-based predictions for competent monitors in terms of the behavioural indexes defined. The Read & Answer software (Vidal-Abarca & Cerdán, 2013) was used to record online each participant's behaviour as well as the product of the task (i.e., their written report).

- A specific monitoring task was designed to allow researchers, in a first approximation, assessing the reliability of predicted causal links between students' processing behaviours and their task products.

In addition, and in the limited scope of this study, we assessed the ability of the extended model to infer the probable cause of monitoring failure when worked-out algebra problem examples are processed. In the long term, the present study would aim at modestly contributing to help teachers infer specific (unobserved) students' behaviours from their observed task outcomes.

Monitoring Worked-out Examples

Worked-out examples, in the form of fully solved problems, are considered important instructional tools in mathematical education (Renkl, 2014), physics education (Docktor & Mestre, 2014; Smith et al., 2010), and chemistry education (Crippen & Earl, 2004; Seery, 2013). Educational approaches based on worked-out examples have been shown to be more effective than other approaches in teaching problem solving to novice students (McLaren et al., 2016) due to its lower cognitive load (Sweller, 2006; van Gog & Rummel, 2010). In teaching their students, teachers often solve and explain a set of interrelated problems (worked-out examples) and then pose new problems that share some characteristics with worked examples (Gómez-Ferragud et al., 2013). While studying worked-out examples, students crucially depend on their own monitoring to detect learning obstacles. A detrimental illusion of understanding and overconfidence could arise if students superficially processed the examples or did not process them till the end (Mihalca et al., 2015). This incomplete processing can happen when the process of understanding the worked-out examples is poorly monitored.

Monitoring has been assessed in calibration studies using worked-out examples (Baars et al., 2014, 2017; Boekaerts & Rozendaal, 2010), problem-solving (Nelson & Fyfe, 2019), or ill-structured problems such as a chess endgame (De Bruin et al., 2005). The reported main cause of students' poor monitoring was the overestimation of their own performance.

In addition to calibration, error detection (Baker, 2002) is also considered a suitable technique to assess students' monitoring (Dehaene, 2018). In an exploratory study on monitoring in mathematics conducted with high school students, different errors were deliberately embedded in different locations of solved examples. These errors sought to hinder students' elaboration of a mental representation of the problematic situation described in the statement, or the translation into algebra of the statement ideas, or the navigation in the problem space (Gómez-Ferragud et al., 2016). In general, students showed poor monitoring, as only 15 percent reported difficulties associated to embedded errors. However, the highest percentage of error detection (25 percent) was observed when the error hindered the algebraic translation process.

A Model for Inconsistency Detection when Processing Worked-out Algebra-Problem Examples

Kintsch (1998) proposed the Construction-Integration (CI) model for text comprehension, based on propositions as meaning units, and processing cycles due to the limitations of the human working memory. From this model, Otero and Kintsch (1992) generated a mechanism explaining mental processes involved in comprehension monitoring when reading texts. They were able to simulate the observed human failure to detect inconsistencies in a text. According to this mechanism, the processing behaviour yielding to the detection of an inconsistency necessarily involves:

1) Mental construction of inconsistent propositions. If a reader processed the information in an erroneous or in a superficial way (paying little attention, for instance), some of these propositions could be poorly constructed or not constructed.

2) Simultaneous processing of inconsistent propositions, i.e. their processing in the same cycle. Propositions must necessarily be compared in the working memory to detect their incompatibility.

3) Both inconsistent propositions must result in being activated at the end of the processing cycle. In this case, difficulties to integrate both propositions in a coherent mental representation could appear and the inconsistency could be detected. If one proposition was suppressed at the end of the processing cycle a student could not be aware of the incompatibility.

If a student detected the inconsistency, then they might initiate observable regulatory actions (Nelson & Narens, 1990) to overcome the detected integration obstacle. Observable actions include seeking for additional information in sources as textbooks or internet, re-reading again and again to be sure the incompatibility detected really exists, asking questions to teacher or peers, or pointing out the detected comprehension obstacle. In the present study, we expected that appropriate regulatory actions would imply:

4) A reader's discrimination of inconsistent from non-conflictive information. This could be observed in the distribution of time processing among units.

Failures Detecting Inconsistencies in Worked-out Algebra-Problem Examples

Otero and Kintsch (1992) simulated failures detecting inconsistencies using a subject's prior knowledge or beliefs as excitatory factors activating one of the contradictory propositions and inhibiting the other one (a kind of priming). In the present study we expected that failures detecting inconsistencies in a worked-out example of algebraic nature could be caused by two main factors different from priming: inappropriate construction of propositions and working memory overloading, and both might be associated with readers' overconfidence.

First, in the case of worked-out algebra-problem examples two different languages, plain language and mathematics, are implied and have to be interconnected. Meaning construction from the statement sentences (plain text) allows a reader to elaborate a Situation Model (SM). The Problem Model (PM) is built from algebraic equations relating important quantities. "The translation process" (Jupri & Drijvers 2016) requires connecting SM with PM: the general and abstract meaning of (some) algebraic expressions has to be particularized or instantiated according with the SM. This can be a source of obstacles for novice students.

Secondly, monitoring the necessary PM-SM connection can overload a reader's working memory because it demands combining two languages while checking for coherence. Information in both languages must be run in the working memory, increasing the probability of overloading. To avoid overloading, a reader could reduce the number of propositions running at once in their working memory. This reduction could decrease the number of links among different propositions. To compensate for this reduction, a reader must generate additional processing cycles to deliberately link text and algebra information units. Otherwise, decreasing the number of links would decrease the probability of inconsistency detection.

Finally, as found in previous studies, a reader's overconfidence can cause careless processing generating both types of obstacles: poor meaning construction and lack of links among propositions.

Method

Participants

Ninety-four Spanish students, 30 women and 64 men, participated in this study. They all were graduates in different STEM specialities: industrial engineering (41), environmental (11), telecommunications (8), electrical and electronics (15), and architecture and civil engineering (19). They were enrolled in a master's degree to become secondary teachers. All of them were informed on the educational purpose and procedures of the research and requested for voluntary participation. Although there was no sampling procedure, these students did not show any special feature compared to those in the corresponding population in Spain.

Participants were expected to have sufficient knowledge and experience in problem solving. Therefore, their potential obstacles in a monitoring-for-understanding task should be mainly due to failures in monitoring and not poor understanding. This should simplify data analysis.

Instruments

Two experimental problems were used in this study. They had different statements, but the same quantities and the same relationships between them (i.e., both had the same algebraic structure). Statements and the full solving procedure were segmented in information units, i.e., complete sentences or equations, as shown in Table 1 (text units were translated from Spanish into English). A 'reversal mistake' (Cooper 1986; González-Calero et al., 2015) was embedded in equation R1, causing this equation to be wrong, i.e., inconsistent with unit S1 (see footnote in Table 1).

Table 1. Information Units in the Experimental Problems

Statement A	Statement B
S1: In a school there are twenty-nine times as many students (A) as teachers (P). S2: There are six hundred people in total. S3: How many teachers are there at this school?	S1: In a chemical dissolution there is twenty-nine times as many water (W) as acid (C), in volume. S2: There are six hundred cubic centimetres in total. S3: How much acid is there in this dissolution?
Resolution R1: $29 * A = P$[1] R2: $A + P = 600$ R3: $29 * (600 - P) = P$ R4: $17400 = 30 * P$ R5: $P = 580$	Resolution R1: $29 * W = C$[1] R2: $W + C = 600$ R3: $29 * (600 - C) = C$ R4: $17400 = 30 * C$ R5: $C = 580$
Sol: Five hundred and eighty teachers	Sol: Five hundred and eighty cubic centimetres of acid

Note. [1]These equations contain the embedded error. The correct ones are: R1': $29 * P = A$; R1': $29 * C = W$.

Data collection was done using the Read & Answer software (Vidal-Abarca & Cerdán, 2013). The software shows information units on the computer screen in a masked, unreadable way, but their ordered sequence is evident. When clicking on a unit, this unit becomes full readable, but the others are (re)masked and thus only one unit can be read at once. Read & Answer allows the readers to read the units in the order they want, and as many times as they need. Pressing a button, participants can move from problem screen to response screen, and the other way round. In the response screen there is a blank space to write the answer to the task. The answer can be modified as many times as needed.

Read & Answer records the full sequence of actions taken by each participant and the time devoted to each action: unmasking information units, moving from one screen to the other, reading the task, or writing/modifying the answer.

In an error detection study, Otero et al. (1992) found participants who detected embedded incoherencies but did not report them. Hence, a brief individual questionnaire was used in the present study to detect participants with inappropriate regulation. In the questionnaire, the embedded error was revealed, and then several questions were asked to know whether or not a) a participant detected the embedded error during the task and b) reported this detection, and c) if the experimental error was detected but was not reported, what was the reason why.

Variables and Measures

According to their written reports (i.e. written responses given to the proposed task), participants were classified as reporters or non-reporters of the embedded error.

Inspired by the assumed theoretical model, a set of indexes were defined to describe participants' monitoring behaviour in the task. These indexes are shown and related to the assumed theoretical model in Table 2.

The indexes were computed as follows:

1. Inconsistent units, S1 and R1, are read at a normal pace or slower. For each participant, normal reading speeds for textual and for algebraic information were computed by averaging their reading visits

Table 2. Indexes Associated to the Process of Detecting and Reporting the Embedded Error

Model components	Indexes
	Detection
1) The mental construction of the two inconsistent propositions, (part of) the meaning of S1 and R1.	1) At least one visit to read each conflicting unit, S1 and R1, must have a reading speed in the normal range or higher (No shallow processing).
2) The simultaneous processing of the inconsistent propositions.	2) The reading sequence has to include one or more consecutive visits to the non-consecutive units S1 and R1 (reading links).
3) The conflicting units result activated at the end of the processing cycle causing integration difficulties.	3) A Slowdown effect, i.e. a significant decrease of the reading speed of R1 in S1-R1 links due to difficulties integrating both units in a coherent mental representation.
	Regulation
4) Discriminating and Focusing on conflicting information. (Regulatory actions due to the participant's awareness of the inconsistency).	4) Higher cumulated times in S1 or in R1 for participants with correct regulatory actions aimed at: a) assuring the contradiction really exists; b) elaborating a written report with explicit mention of the contradiction between R1 and S1.

to textual or to algebraic units. 'Normal' reading pace refers to the usual, standard speed of an individual when they do a complete processing of information units, instead of skimming or scanning actions.

The measures used were mean time per word and mean time per symbol (and standard deviations, *SD*), which are inverse to reading speed expressed in words, or symbols, per minute. In algebraic units, every symbol was computed: numbers (one symbol per quantity), letters representing quantities, operations, parentheses, and the equal sign. For instance, unit R3 has 9 symbols: 29, ∗, (, 600, - , P ,) , = , P.

For text units, and only when computing individual normal reading speed, values lower than 100 ms per word were discarded for computing because they are usually considered as incompatible with meaning access (Sereno et al., 1998). Expert readers show speeds of 250-400 words per minute in reading for understanding, corresponding to 150-240 ms/word (Rayner et al., 2016). Thus, 100 ms/word is a conservative lower limit.

For algebra units Jansen et al. (2007) computed mean values of 240 ms for fixation times when experts read algebraic expressions. Andrà et al., (2015) obtained mean values of 190-250 ms for fixation times in a sample of students. When means and standard deviations in these studies are considered, it seems that 100 ms/symbol is also a very conservative lower limit for meaning access in algebra reading as well. Visits with times lower than 100 ms/symbol were excluded when computing individual reading speed of algebra units.

For each participant, the values (in ms/word or ms/symbol) of the remaining reading visits were averaged and mean and standard deviation were obtained. Values in the range [mean - *SD*, mean + *SD*] were considered 'normal' for each participant in the proposed task.

2. Existence of explicit reading links between S1 and R1. Due to working memory limitations and the overloading translation process, we expected that most participants would need to make a reading S1-R1 link, i.e., a consecutive reading visit of non-consecutive units S1 and R1, to simultaneously process these units in the same cycle.

3. Existence of slowdowns when reading inconsistent units. A slowdown was computed for each participant when time per word or per symbol in a reading visit was higher than their normal upper limit, mean + *SD.* According to the CI model (Kintsch, 1998), when incompatible propositions resulted activated at the end of the same processing cycle there would be processing difficulties in the integration phase. Integration difficulties could be observed as reading slowdowns. Van der Schoot et al. (2012) observed differences in reading speed equivalent to 125-150 ms/word when they compared coherent versus incoherent sentence conditions. For the present task, and due to the location of the embedded error in R1, the slowdown due to integration obstacles was expected to be mainly observed when processing R1 in a S1-R1 link.

In addition, and because of regulatory actions done (i.e., re-readings to be sure of detection and re-readings to elaborate the report), reporters were expected to make a greater processing effort in S1 and R1 than non-reporters. In the present study, the "processing effort" for a unit u_j was defined as the ratio between the reading time cumulated in u_j, and the time needed to read u_j once at normal speed. Therefore, individual differences in reading speed were taken into account to obtain comparable and unbiased values.

Predictions

The model assumed suggested the following predictions for the designed task:

1. Successful monitors should show slowdowns in R1 (which includes the error) in some S1-R1 link (see Table 2).

2. Reporting the detected error in R1 should be associated with a processing discrimination of S1 and R1 from the remaining non-conflicting units (index 4 in Table 2).

In addition, and a little beyond predictions of the model:

3. The defined set of behavioral indices should differentiate the processing of competent monitors from that of poor monitors.

Procedure

Each participant completed the task in one of the experimental problems, randomly assigned to the computers in a specific room. The instructions explicitly included the purpose of the study, the task, and a brief explanation of the Read & Answer software. The task proposed to the students in the instructions was: "Judge the understandability of a completely solved algebra problem and report in writing anything you think is incorrect, nonsensical, inconsistent, or unintelligible. Try to clarify the location and reason for each reported issue, if any." We tried to avoid a 'seek for the error' task that can skew student performance due to an increased awareness.

A practice task was conducted using a worked-out example, similar to the experimental ones. This practice was useful to train participants in the use of the Read & Answer software and to clarify the task. Next, participants completed the experimental task at their own pace without any additional help. The task lasted less than 30 minutes. Immediately after this task, participants individually completed the questionnaire in a separated room. Responses to the questionnaires were used to increase the reliability of the classification of participants as reporters or non-reporters: some participants could show wrong regulation, i.e., they detected the embedded error but did not mention it.

Results

Task Outcomes: Participants' Reports

Forty-seven out of 94 participants (50.0 percent) correctly reported the embedded error as located in R1: "*La primera ecuación de la*

Table 3. Percentages (and Number) of Reporters and Non-reporters Satisfying Behavioural Indices

Model-based Behavioural Indexes	Reporters	Non-reporters
Normal speed reading S1 and R1 (Index 1)	100 (47)	100 (47)
Links S1-R1 or R1-S1 (Index 2)	95.7 (45)	36.2 (17)
Slowdown in S1 or in R1 associated to links S1-R1 or R1-S1(Index 3)	95.7 (45)	8.5 (4)
Processing efforts on S1 and R1 higher than in other text or algebra units (Index 4)	76.6 (36)	29.8 (14)
Simultaneous satisfaction of indexes 1-4	76.6 (36)	6.4 (3)

resolución está mal. La correcta sería A = 29P. Por tanto, la solución final está mal" ["The first equation in the solution is incorrect. The correct one would be A = 29P. Therefore, the final solution is incorrect"] (student #66 AC); "*El problema está mal resuelto (…). Sería al revés, es decir, A = 29C, por lo que C = 20 ml*" ["The problem is poorly resolved ... It would be the other way around, that is, A = 29C, so that C = 20 ml"] (#86 AMM).

Forty-seven students did not mention the error embedded in R1 and were called non-reporters. Some of them did not report any obstacle: "*No he encontrado nada incorrecto, y creo que el enunciado del problema se entiende perfectamente*" ["I have not found anything wrong, and I think the problem is perfectly understood"] (#46 PT).

Others reported obstacles different from the embedded error (spurious obstacles): "*Para mi hay 2 obstáculos: el primero es leer los números en letras; el segundo es que faltan imágenes*" ["For me there are 2 obstacles: 1st, reading the numbers as letters; 2nd, missing images"] (#61 Pr22).

We also analysed the responses given by non-reporters to the questionnaire and we did not find any participant detecting and non-reporting the embedded error. Here on, non-reporters were considered non-detectors of the embedded error.

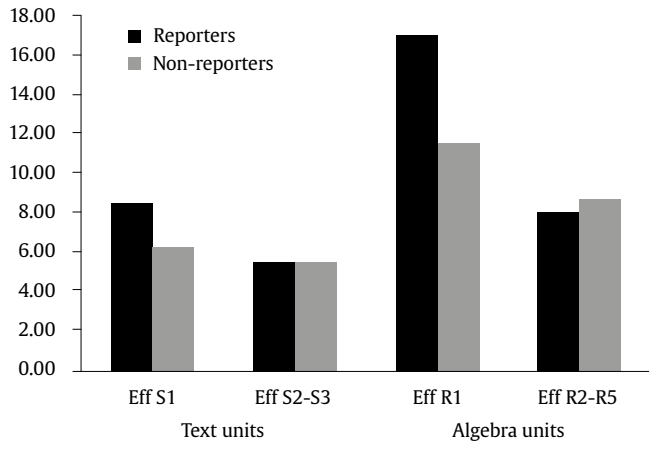

Figure 1. Processing Effort in the Conflicting Units S1and R1 for Reporters and Non-reporters, Compared to the Effort Made in the Remaining Textual or Algebraic Units.

Behavioural Indexes for Reporters and Non-reporters

Table 3 shows the percentages of reporters and non-reporters who satisfied defined behavioural indexes.

All the participants, reporters and non-reporters, read S1 and R1 at a normal pace at least once (index 1 in Table 2). Thus, fast reading was not the cause of monitoring difficulties. Reading links between S1 and R1 (index 2 in Table 2) were made at least once by 95.7 percent of reporters and 36.2 percent of non-reporters. Slowdowns in R1 in S1-R1 links (index 3) were observed in 95.7 percent of reporters, but only in 8.5 percent of non-reporters. Therefore, index 3 showed a high capacity to discriminate reporters from non-reporters.

As the distribution of the processing effort concerns (index 4), a percentage of 76.6 percent of reporters made higher effort in both units, R1 and S1, than in the corresponding statement or algebra units averaged. This only happened in 29.8 percent of non-reporters. Therefore, index 4 did not show a high discriminating capacity. However, when mean values for reporters or non-reporters were considered, some differences appeared in the distribution of the processing efforts. Figure 1 shows these mean values.

Predicting Correct or Poor Reporting from the Behavioural Indexes

In order to evaluate the power of the set of defined indexes to correctly predict a correct reporting or not, a binary logistic regression was conducted for the dependent variable (reporters/non-reporters) taking the behavioural indexes as predictors: the existence or not, and the number of explicit S1-R1 links, the existence or not of slowdowns in R1 in S1-R1 links, the existence or not of slowdowns in S1 in S1-R1 links, the efforts made processing units S1, R1, S2-S3 (two units averaged), Sol, and R2-R5 (four units averaged). The efforts in the different units were transformed using the log function. In this way, these variables turned to be normal (K-S, $p > .200$ in all cases).

When considered alone, these predictors were significant except the effort in S2-S3 and the effort in R2-R5. When a backward stepwise procedure was run, some important predictors were redundant and were left out (existence and number of S1-R1 links, effort in R1). Only the existence or not of slowdowns in S1 in S1-R1 links (S1SlowD in S1-R1), the existence or not of slowdowns in R1 in S1-R1 links (R1SlowD in S1-R1), and the effort in S1 (EffS1) remained in the equation. The model significantly fit the data: omnibus test: $\chi^2(3) = 105.83$, $p < .001$. Together, these three variables explained a large percentage of the variance of the dependent variable (-2 log likelihood = 24.48, Cox & Snell $R^2 = .68$, Nagelkerke $R^2 = .90$) and a global 94.7 percent of cases were correctly classified.

Table 4 shows the relevant data of the binary logistic regression.

The existence or not of slowdowns in R1 in S1-R1 links was the most important predictor (Nagelkerke $R^2 = .84$), and the existence or not of slowdowns in S1 in S1-R1 links was the less important one (adding 2 percent points to R^2).

Table 4. Main Data of the Logistic Binary Regression

Predictors	*B*	*SE*	Wald's test	*df*	*p*	Exp(*B*) (odds ratio)
Constant	-2.209	1.110	3.960	1	< .05	0.11
S1SlowD in S1-R1	2.618	1.361	3.704	1	.05	13.71
R1SlowD in S1-R1	9.453	2.568	13.547	1	< .001	12,747.42
EffS1	-5.958	2.230	7.138	1	< .01	0.00

The effort in S1 obtained a negative coefficient in the regression. However, since the effort in S1 has a positive coefficient when considered alone, this is merely a mathematical effect to correct an overprediction obtained with the other two variables.

Table 5 shows the observed and predicted result in the task (reporter/non-reporter).

Table 5. Observed versus Predicted Number of Reporters and Non-reporters after the Binary Logistic Regression

	Predicted		
Observed	Reporters	Non-reporters	% Correct
Reporters	45	2	95.7
Non-reporters	3	44	93.6
Total	48	46	
Overall % correct			94.7

As can be observed in Table 5, the probability of being a reporter when the predicted behaviour was performed (according to the predictors in the resulting logit regression), was .94. The probability of being a non-reporter when the expected behaviour was not performed was .96. Conversely, the observed product of the task (correct reporting of the embedded error/failed reporting) could be used to infer a participant's probable behaviour when processing the information provided. The probability of performing the predicted processing behaviour when a correct report was observed was .96, since behavioural indexes were satisfied by 95.7 percent of the (observed) reporters. Only two anomalous reporters did not behave as predicted by model-based indexes. These cases will be analysed in the Discussion section below. In addition, the probability of any of the predicted processing actions was not performed when a failed reporting was observed reached .94, since 93.6 percent of the observed non-informants did not meet at least one of the behavioral indexes. Finally, the accuracy of the prediction made by the binary logistic regression was .95.

Discussion

Predictions for reporters were fulfilled in a high proportion of cases. In addition, the results obtained suggested different behavioural patterns for competent or poor monitors in the task, as described by behavioural indexes.

Model-Predicted Behaviour for Reporters

All the reporters but two (95.7 percent) fitted the predictions of the theoretical model, according to defined behavioural indexes. All of them processed R1 and S1 in a non-superficial way, and 95.7 percent showed a slowdown in R1 in some S1-R1 link. Longer times when reading conflicting units, compared to non-conflicting units, have also been observed in previous experiments (Van der Schoot et al. 2012). Using refutation texts Mason et al. (2019) observed delays in response times attributed to participants' inhibition of a prior knowledge inconsistent with the text.

However, two anomalous reporters appeared. One of the unfitting reporters (participant #94 BdT) did not explicitly link R1 to S1 in his reading sequence (index 2). He detected the absurd numerical result from his prior knowledge about usual schools, but he probably assumed that it was due to a letter-switch mistake and not to a translation error. Therefore, he initiated a backwards processing from R4 to R1 searching for the first time A and P were reversed. In this way he did not need to re-read S1, and then no links appeared between S1 and R1. The other anomalous reporter (participant #58 LF) did not show any slowdown in R1 associated to S1-R1 links (index 3), but she satisfied indexes 2 and 4 and made slowdowns in R1 out of S1-R1 links. This participant was probably able to carry part of the meaning of S1 ('there is more students than teachers in a school'), well-known information, over several cycles to link it with R1 in an implicit, unobservable manner.

Model-based Behavioural Indexes for Non-reporters

In addition to the model-based predictions for reporters, 91.5 percent of non-reporters (43 participants out of 47) did not show any reading slowdown in R1 in S1-R links (index 3). Only 8.5 percent of non-reporters (n = 4) showed at least a slowdown in R1 in a S1-R1 link. This suggests that these students had difficulties integrating S1 and R1. However, they reported spurious errors in the monitoring task attributed to R1. Hence, they probably built some inappropriate meaning for R1 (poor construction of propositions), causing integration difficulties different from the predicted one.

Among non-reporters without any slowdown in R1 in S1-R1 links, 63.8 percent (n = 30) did not make any S1-R1 link. This alerts to a possible lack of metacognitive skills in adult students. Long time ago Vosniadou et al. (1988) observed that most children were unable to detect inconsistencies in a text because they had difficulties to re-activate preceding propositions to be processed with other subsequent contradictory propositions. When the two contradictory propositions were simultaneously (re)activated, children detected the contradiction. In terms of the CI model (Kintsch, 1998), this suggested that children's difficulties appeared in the construction phase, but not in the integration phase once the construction of the propositions was appropriate. In our study, if those non-reporters had processed S1 and R1 together making deliberated S1-R1 reading links, they probably would have detected the contradiction. As the translation of the units into the same language is an overloading process, carrying previously read information over several cycles can be difficult. Thus, the absence of explicit reading links between S1 and R1 could cause monitoring failures in these students.

Changes in the reading sequence, as reading back to link distant pieces of information, have also been observed using eye tracking techniques when people process inconsistent or difficult information (Hyönä et al., 2003). When comprehension fails, readers try to overcome comprehension obstacles by re-reading previous text segments to re-establish coherence. Rinck et al. (2003) used texts with embedded temporal inconsistencies in separated sentences. Readers could report any comprehension obstacle they found. Specific differences between reporters and non-reporters were obtained in (a) the number of reading regressions linking both contradictory propositions in the text and (b) the reading speed in the 2nd contradictory proposition after the regressions. When regressions to re-read were impeded, readers had additional difficulties and comprehension was greatly compromised (Schotter et al., 2014), even in normal reading (Rayner et al., 2016).

The remaining 27.7 percent (n = 13) of non-reporters made S1-R1 links without any slowdown in R1. Thus, they seemed to integrate S1 and R1 without special difficulties. According to the assumed theoretical model, this lack of integration difficulties could be due to a suppression of one of the two inconsistent propositions (for instance, the part of R1 meaning that there would be more teachers than students at that school) or to the incorrect or poor construction of any of the inconsistent propositions (for instance, when a student accepted the reversal error in R1 as correct). A detailed analysis showed that 10.7% (5 students) reported spurious errors caused by poor construction of meaning; 4.2% (2 students) agreed with the offered resolution accepting the reversal error in R1 as correct, probably due to poor construction of propositions as well; 12.8% of non-reporters (6 students) reported that something was wrong in the resolution without any mention of the error in R1. These participants

seemed unable to find the embedded error in R1 despite searching for it, probably due to the suppression or one of the inconsistent propositions (for instance, the part of R1 meaning that there were more teachers than students at that school).

In sum, and according to the theoretical model, 23.4 percent of non-reporters seemed to show poor or wrong construction of propositions (mainly R1), 63.8 percent achieved low levels of coherence in their mental representations mainly due to lack of reading links in an overloading task, and 12.8 percent probably suppressed one of the inconsistent propositions when they processed them together.

In addition, reporters seemed to focus their processing effort on the conflicting units more than non-detectors in average. However, index 4 by itself did not properly discriminate regulatory actions made by reporters or non-detectors. Regulation can be a complex process (Chinn & Brewer, 1993), and additional work is needed to validate well-articulated models for regulation.

A specific combination of behavioural indexes 4 and 3, given by a binary logistic regression, correctly discriminated reporters from non-reporters in a 94.7 percent of the cases. The observed competent monitoring could be correctly associated to the model predicted processing behaviour in 95.7 percent of the reporters, and the poor monitoring could be correctly associated to a failure in the expected behaviour in 93.6 percent of non-reporters.

Additional Data

In the experimental task used in the present study, participants showed global poorer monitoring than expected according to their education level (postgraduate engineers). This poor monitoring ability has been found in academic problem-solving (Baars Van Gog et al., 2014, 2017; García et al., 2016; Nelson & Fyfe, 2019). In these studies participants showed poor calibration due to extra confidence although they were strategic in their help-seeking decisions. Low levels of monitoring in secondary students were also reported in a reading-for-answering monitoring task by Sanjosé et al. (2010). In their experiments only about 46 percent of participants were able to reject a 'matching words' strategy, i.e., selecting inappropriate information made of the same words as present in questions asked, to elaborate incorrect answers.

In addition to overconfidence, the low level of achievement in the present study could also be due to the difficulty of the monitoring task, which was harder than expected for the participants. This conjecture is supported by the low mean reading speed obtained in the present study for textual units, 106 words per minute, compared to the values of 200-400 words per minute obtained in other studies for adult normal reading (Rayner et al., 2016).

Regarding algebra units, only few studies reported reading speeds. Andrà et al., (2015) obtained a mean value of about 240 ms/symbol for students' average fixation times reading algebraic expressions. This mean value is very similar to the mean value obtained if experts' data obtained by Jansen et al. (2007) is applied to units R1-R5 of the experimental problem used in the present study. However, in the present experiment participants' mean reading speed was clearly slower, 536 ms/symbol (SD = 335), or 112 algebra symbols per minute. The great difference with previous studies can be explained by the different tasks to be performed: reading for comprehension and checking for coherence in the present study, or only reading algebraic expressions in previous studies.

Limitations

The present study has limitations caused by different factors. First, varied monitoring tasks and experimental working-out examples should be used in future studies. Second, participants had algebra knowledge and problem-solving expertise higher than secondary students. These last students would probably show cognitive in addition to metacognitive difficulties when doing a task as the proposed here. Therefore, the present study should be conducted with secondary students to increase reliability of results.

Conclusions

The theoretical model used in the present study—an extension of Otero and Kintsch's (1992) mechanism of inconsistency detection to worked-out examples of algebraic nature—suggested specific predictions for reporters' (competent monitors') behaviours in the monitoring task. These behaviours were characterized by a set of process-based indexes. The predicted behaviour was observed in a high percentage of competent monitors, with only two exceptions.

Beyond the model prediction, the defined process-based indexes were also tested in their power to correctly discriminate competent from poor monitors' behaviour. A binary logistic regression for the reporters/non-reporters variable correctly classified a very high percentage of cases from a specific combination of behavioural indexes. Probabilities for the association of the observed outcome in the task (correct reporting or not) with the predicted processing behaviour reached high values. Therefore, if these results were replicated in a variety of situations and worked-out examples, a student's observed competency or poor monitoring in this task should be associated with a specific processing behaviour with high reliability and teachers could help them in a more specific, accurate way. For instance, teachers could assume with a low risk that observed students' poor monitoring could be mainly due to an insufficient effort devoted to attain high global coherence in their mental representations or to great difficulties to construct some important propositions (for instance, the meaning of R1 in the present study). Second, teachers must expect students' reading speeds of worked-out examples as low as those reported here (about 106 words and 112 symbols per minute, i.e., about three times slower than students' reading speeds for usual texts) when studying worked-out examples carefully. Faster reading (or fast explanations!) will probably be associated to shallow processing.

Conflict of Interest

The authors of this article declare no conflict of interest.

Acknowledgement

Authors thank members of "ERI-Lectura", University of Valencia, Spain, for providing them with the Read & Answer software.

References

Andrà, C., Lindström, P., Arzarello, F., Holmqvist, K., Robutti, O., & Sabena, C. (2015). Reading mathematics representations: An eyetracking study. *International Journal of Science and Mathematics Education, 13*(2), 237-259. https://doi.org/10.1007/s10763-013-9484-y

Baars, M., Van Gog, T., de Bruin, A., & Paas, F. (2014). Effects of problem solving after worked example study on primary schoolchildren's monitoring accuracy. *Applied Cognitive Psychology, 28*(3), 382-391. https://doi.org/10.1002/acp.3008

Baars, M., Van Gog, T., de Bruin, A., & Paas, F. (2017). Effects of problem solving after worked example study on secondary school children's monitoring accuracy. *Educational Psychology, 37*(7), 810-834. https://doi.org/10.1080/01443410.2016.1150419

Baker, L. (2002). Metacognition in comprehension instruction. In C. Collins & S. R. Parris (Eds.), *Comprehension instruction: Research-based best practices* (pp. 65-79). Guilford Press.

Baten, E., Praet, M., & Desoete, A. (2017). The relevance and efficacy of metacognition for instructional design in the domain of mathematics. *ZDM Mathematics Education, 49*(4), 613-623. https://doi.org/10.1007/s11858-017-0851-y

Boekaerts, M., & Rozendaal, J. S. (2010). Using multiple calibration indices in order to capture the complex picture of what affects students' accuracy of feeling of confidence. *Learning and Instruction, 20*(5), 372-382. https://doi.org/10.1016/j.learninstruc.2009.03.002

Cerdán, R., Gilabert, R., & Vidal-Abarca, E. (2011). Selecting information to answer questions: Strategic individual differences when searching texts. *Learning and Individual Differences, 21*(2), 201-205. https://doi.org/10.1016/j.lindif.2010.11.007

Chinn, C. A., & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. *Review of Educational Research, 63*(1), 1-49. https://doi.org/10.3102/00346543063001001

Cooper, M. (1986). The dependence of multiplicative reversal on equation format. *Journal of Mathematical Behaviour, 5*(2), 115-120.

Cornoldi, C., Carretti, B., Drusi, S., & Tencati, C. (2015). Improving problem solving in primary school students: The effect of a training programme focusing on metacognition and working memory. *British Journal of Educational Psychology, 85*(3), 424-439. https://doi.org/10.1111/bjep.12083

Cox, M. T., Alavi, Z., Dannenhauer, D., Eyorokon, V., Muñoz-Avila, H., & Perlis, D. (2016). MIDCA: A metacognitive, integrated dual-cycle architecture for self-regulated autonomy. *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence* (pp. 3712-3718). AAAI Press.

Crippen, K. J., & Earl, B. L. (2004). Considering the efficacy of web-based worked examples in introductory chemistry. *Journal of Computers in Mathematics and Science Teaching, 23*(2), 151-167.

De Bruin, A. B., Rikers, R. M., & Schmidt, H. G. (2005). Monitoring accuracy and self-regulation when learning to play a chess endgame. *Applied Cognitive Psychology, 19*(2), 167-181. https://doi.org/10.1002/acp.1109

Dehaene, S. (2018). The error-related negativity, self-monitoring, and consciousness. *Perspectives on Psychological Science, 13*(2), 161-165. https://doi.org/10.1177/1745691618754502

Desoete, A., & De Craene, B. (2019). Metacognition and mathematics education: An overview. *ZDM Mathematics Education. 51*(4), 565-575. https://doi.org/10.1007/s11858-019-01060-w

Docktor, J. L., & Mestre, J. P. (2014). Synthesis of discipline-based education research in physics. *Physical Review Special Topics-Physics Education Research, 10*(2), 020119. https://doi.org/10.1103/PhysRevSTPER.10.020119

García, T., Cueli, M., Rodríguez, C., Krawec, J., & González-Castro, P. (2015). Conocimiento y habilidades metacognitivas en estudiantes con un enfoque profundo de aprendizaje. Evidencias en la resolución de. *Revista de Psicodidáctica, 20*(2), 209-226. https://doi.org/10.1387/RevPsicodidact.13060

García, T., Rodríguez, C., González-Castro, P., González-Pienda, J. A., & Torrance, M. (2016). Elementary students' metacognitive processes and post-performance calibration on mathematical problem-solving tasks. *Metacognition and Learning, 11*(2), 139-170. https://doi.org/10.1007/s11409-015-9139-1

González-Calero, J. A., Arnau, D., & Laserna-Belenguer, B. (2015). Influence of additive and multiplicative structure and direction of comparison on the reversal error. *Educational Studies in Mathematics, 89*(1), 133-147. http://doi.org/10.1007/s10649-015-9596-0

Gómez-Ferragud, C. B., Sanjosé V., & Solaz-Portolés, J. J. (2016). Estudios sobre comprensión y control de la comprensión en resolución de problemas académicos. *Revista de Enseñanza de la Física, 28*(1), 21-35.

Gómez-Ferragud, C. B., Solaz-Portolés, J. J., & Sanjosé, V. (2013). Analogy construction and success in mathematics and science problem-solving: A study with secondary students. *Revista de Psicodidáctica, 18*(1), 81-111.

Hartman, H. J. (2001). Developing students' metacognitive knowledge and skills. In H. Hartman (Ed.), *Metacognition in learning and instruction* (pp. 33-68). Kluwer.

Hyönä, J., Lorch, Jr. , R. F., & Rinck, M. (2003). Eye movement measures to study global text processing. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 313-334). Elsevier.

Jansen A. R., Marriott K., & Yelland G. W. (2007) Parsing of algebraic expressions by experienced users of mathematics. *European Journal of Cognitive Psychology, 19*(2), 286-320. https://doi.org/10.1080/09541440600709955

Jupri, A., & Drijvers, P. H. M. (2016). Student difficulties in mathematizing word problems in algebra. *Eurasia Journal of Mathematics, Science and Technology Education, 12*(9), 2481-2502. https://doi.org/10.12973/eurasia.2016.1299a

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

Koçak, R., & Boyaci, M. (2010). The predictive role of basic ability levels and metacognitive strategies of students on their academic success. *Procedia-Social and Behavioural Sciences, 2*(2), 767-772. https://doi.org/10.1016/j.sbspro.2010.03.099

Lucangeli, D., Penna, M. P., Fastame, M. C., Pedron, M., Porru, A., & Duca, V. (2019). Metacognition and errors: The impact of self-regulatory trainings in children with specific learning disabilities. *ZDM Mathematics Education, 51*(4). https://doi.org/10.1007/s11858-019-01044-w

Mañá, A., Vidal-Abarca, E., & Salmerón, L. (2017). Effect of delay on search decisions in a task-oriented reading environment. *Metacognition and Learning, 12*(1), 113-130. https://doi.org/10.1007/s11409-016-9162-x

Mason, L., Zaccoletti, S., Carretti, B., Scrimin, S., & Diakidoy, I. A. N. (2019). The role of inhibition in conceptual learning from refutation and standard expository texts. *International Journal of Science and Mathematics Education, 17*(3), 483-501. https://doi.org/10.1007/s10763-017-9874-7

McLaren, B. M., van Gog, T., Ganoe, C., Karabinos, M., & Yaron, D. (2016). The efficiency of worked examples compared to erroneous examples, tutored problem solving, and problem solving in computer-based learning environments. *Computers in Human Behaviour, 55*, 87-99. https://doi.org/10.1016/j.chb.2015.08.038

Mihalca, L., Mengelkamp, C., Schnotz, W., & Paas, F. (2015). Completion problems can reduce the illusions of understanding in a computer-based learning environment on genetics. *Contemporary Educational Psychology, 41*, 157-171. https://doi.org/10.1016/j.cedpsych.2015.01.001

Modestou, M., & Gagatsis, A. (2010). Cognitive and metacognitive aspects of proportional reasoning. *Mathematical Thinking and Learning, 12*(1), 36-53. https://doi.org/10.1080/10986060903465822

Morales, J., Lau, H., & Fleming, S. M. (2018). Domain-general and domain-specific patterns of activity supporting metacognition in human prefrontal cortex. *Journal of Neuroscience, 38*(14), 3534-3546. https://doi.org/10.1523/JNEUROSCI.2360-17.2018

Nelson, L. J., & Fyfe, E. R. (2019). Metacognitive monitoring and help-seeking decisions on mathematical equivalence problems. *Metacognition and Learning, 14*(2), 167-187. https://doi.org/10.1007/s11409-019-09203-w

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In H. B. Gordon (Ed.), *Psychology of learning and motivation, 26* (pp. 125-173). Academic Press.

Otero, J. (2009). Question generation and anomaly detection in texts. In *Handbook of metacognition in education* (pp. 59-71). Routledge.

Otero, J., Campanario, J. M., & Hopkins, K. D. (1992). The relationship between academic achievement and metacognitive comprehension monitoring ability of Spanish secondary school students. *Educational and Psychological Measurement, 52*(2), 419-430. https://doi.org/10.1177/0013164492052002017

Otero, J., & Graesser, A. C. (2001). PREG: Elements of a model of question asking. *Cognition and Instruction, 19*(2), 143-175. https://doi.org/10.1207/S1532690XCI1902_01

Otero, J., Ishiwa, K., & Sanjosé, V. (2008). Readers' questioning: Some hints for automated question generation. *Proceedings of the Workshop on the Question Generation Shared Task and Evaluation Challenge*. Arlington, VA.

Otero, J., & Kintsch, W. (1992). Failures to detect contradiction in a text: What readers believe versus what they read. *Psychological Science, 3*(4), 229-235. https://doi.org/10.1111/j.1467-9280.1992.tb00034.x

Özsoy, G., & Ataman, A. (2017). The effect of metacognitive strategy training on mathematical problem-solving achievement. *International Electronic Journal of Elementary Education, 1*(2), 67-82.

Qiu L., Su J., Ni, Y., Bai, Y., Zhang, X., Li, X., & Wang, X. (2018). The neural system of metacognition accompanying decision-making in the prefrontal cortex. *PLoS Biology 16*(4), e2004037. https://doi.org/10.1371/journal.pbio.2004037

Rayner, K., Schotter, E. R., Masson M. E. J., Potter M. C., & Treiman R. (2016). So much to read, so little time: How do we read, and can speed reading help? *Psychological Science in the Public Interest, 17*(1) 4-34. https://doi.org/10.1177/1529100615623267

Renkl, A. (2014). Toward an instructionally oriented theory of example-based learning. *Cognitive Science, 38*(1), 1-37. https://doi.org/10.1111/cogs.12086

Rinck, M., Gámez, E., Díaz, J. M., & De Vega, M. (2003). Processing of temporal information: Evidence from eye movements. *Memory & Cognition, 31*(1), 77-86. https://doi.org/10.3758/BF03196084

Salmerón, L., Naumann, J., García, V., & Fajardo, I. (2017). Scanning and deep processing of information in hypertext: An eye tracking and cued retrospective think-aloud study. *Journal of Computer Assisted Learning, 33*(3), 222-233. https://doi.org/10.1111/jcal.12152

Sánchez, E., & García-Rodicio, H. (2013). Using online measures to determine how learners process instructional explanations. *Learning and Instruction, 26*(1), 1-11. https://doi.org/10.1016/j.learninstruc.2012.12.003

Sanjosé, V., Fernández, J. J., & Vidal-Abarca, E. (2010). Importancia de las destrezas de procesamiento de la información en la comprensión de textos científicos. *Infancia y Aprendizaje, 33*(4), 529-541. https://doi.org/10.1174/021037010793139581

Sanjosé, V., & Torres, T. (2014). Questions generated on non-textual information: An empirical validation of the obstacle-goal model applied to the comprehension of scientific experimental devices. *Universitas Psychologica, 13*(1), 357-368.

Schotter, E. R., Tran, R., & Rayner, K. (2014). Don't believe what you read (only once): Comprehension is supported by regressions during reading. *Psychological Science, 25*(6), 1218-1226. https://doi.org/10.1177/0956797614531148

Seery, M. K. (2013). Harnessing technology in chemistry education. *New Directions in the Teaching of Physical Sciences, 9*(1), 77-86. https://doi.org/10.11120/ndir.2013.00002

Sereno, S. C., Rayner, K., & Posner, M. I. (1998). Establishing a timeline of word recognition: Evidence from eye movements and event-related potentials. *Neuroreport, 9*(10), 2195-220. https://doi.org/2195-2200.10.1097/00001756-199807130-00009

Smith, A. D., Mestre, J. P., & Ross, B. H. (2010). Eye-gaze patterns as students study worked-out examples in mechanics. *Physical Review Special*

Topics-Physics Education Research, 6(2), 020118. https://doi.org/10.1103/PhysRevSTPER.6.020118

Sun, R., Zhang, X., & Mathews, R. (2006). Modeling meta-cognition in a cognitive architecture. *Cognitive Systems Research, 7*(4), 327-338. https://doi.org/10.1016/j.cogsys.2005.09.001

Sweller, J. (2006). The worked example effect and human cognition. *Learning and Instruction, 16*(2), 165-169. https://doi.org/10.1016/j.learninstruc.2006.02.005

Throndsen, I. (2011). Self-regulated learning of basic arithmetic skills: A longitudinal study. *British Journal of Educational Psychology, 81*(4), 558-578. https://doi.org/10.1348/2044-8279.002008

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. https://doi.org/10.1007/s11145-011-9337-4

Van Gog, T., & Rummel, N. (2010). Example-based learning: Integrating cognitive and social-cognitive research perspectives. *Educational Psychology Review, 22*(2), 155-174. https://doi.org/10.1007/s10648-010-9134-7

Vidal-Abarca, E., & Cerdán, R. (2013). Read & answer: An application to study task-oriented reading situations. *Information Design Journal, 20*(1), 70-78. https://doi.org/10.1075/idj.20.1.07vid

Vosniadou, S., Pearson, P. D., & Rogers, T. (1988). What causes children's failures to detect inconsistencies in text? Representation versus comparison difficulties. *Journal of Educational Psychology, 80*(1), 27-39. https://doi.org/10.1037/0022-0663.80.1.27

Wang, M. C., Haertel, G. D., & Walberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research, 63*(3), 249-294. https://doi.org/10.2307/1170546

Yeung, N., & Summerfield, C. (2014). Shared mechanisms for confidence judgements and error detection in human decision making. In S. Fleming & C. Frith (Eds.), *The cognitive neuroscience of metacognition* (pp. 147-167). Springer

Young, A., & Fry, J. D. (2008). Metacognitive awareness and academic achievement in college students. *Journal of the Scholarship of Teaching and Learning, 8*(2), 1-10.