Metacognitive Judgments as an Emerging Research Trend. A Conceptual Review*

(English Version)

Los juicios metacognitivos como tendencia emergente de investigación. Una revisión conceptual

Julgamentos metacognitivos como uma tendência emergente de pesquisa. Uma revisão conceitual

Received 22 March, 2022. Accepted 31 April, 2022.

Adriana Julieth Olaya Torres**
https://orcid.org/0000-0002-1376-9383
Chile

Diana Marcela Montoya Londoño***
https://orcid.org/0000-0001-8007-0102
Colombia

Antonio P. Gutierrez de Blume****
https://orcid.org/0000-0001-6809-1728
United States

Aníbal Puente Ferreras*****
https://orcid.org/0000-0003-3994-0553
Chile

› *To cite this article:*
Olaya Torres, A. J., Montoya Londoño, D. M., Gutierrez, A. P. & Puente, A. (2023). Metacognitive Judgments as an Emerging Research Trend. A Conceptual Review.
Ánfora, *30*(54). 254-281.
https://doi.org/10.30854/anfv30.n54.2023.910
Universidad Autónoma de Manizales. L-ISSN 0121-6538.
E-ISSN 2248-6941.

* None of the authors have any conflict of interest to report.

** PhD student in Developmental Sciences and Psychopathology at Universidad del Desarrollo, Santiago, Chile. aolayat@udd.cl

*** PhD in Cognitive Sciences. University of Manizales, University of Caldas, Colombia. diana.montoya@ucaldas.edu.co

**** PhD in Educational Psychology, Foundations. Georgia Southern University, Department of Curriculum, Foundations, and Reading, United States. agutierrez@georgiasouthern.edu

***** PhD in Education and Society, Master in Reading Comprehension and Text Production. Universidad Andrés Bello, Viña del Mar, Chile. anibal.puente@unab.cl

Abstract

Objective: The purpose of the current article is to conduct a conceptual revision of the term "metacognitive judgment." **Methodology**: For the thematic review of the literature written by some of the most important authors in the field, 55 archival sources were taken into consideration. These sources were examined through the following stages: introduction, reflection, and establishment of conclusions. **Results**: It is established that metacognitive judgments are seen as the construct that brings together the students' beliefs about what they know and do not know, as well as their control and regulation over their learning. **Conclusion**: The study of metacognitive judgments is put out as an alternative to aid students in self-regulating their learning so they can become gradually more accurate in assessing their performance.

Keywords: metacognitive judgments; calibration; accuracy; confidence level; performance; metacognitive monitoring.

Resumen

Objetivo: en el presente artículo se tuvo como objetivo realizar una revisión conceptual del constructo 'juicio metacognitivo'. **Metodología**: para la revisión temática de la literatura producida por algunos de los autores más relevantes del campo se consideraron 55 fuentes documentales que fueron analizadas mediante las siguientes etapas: introducción, reflexión y establecimiento de conclusiones. **Resultados**: se establece que los juicios metacognitivos son entendidos como el constructo que agrupa el conjunto de creencias que tienen los estudiantes acerca de lo que saben, y no saben; y también, respecto a cómo controlan y regulan su aprendizaje. **Conclusión**: el estudio de los juicios metacognitivos se presenta como una alternativa para favorecer el proceso de autorregulación del aprendizaje de los estudiantes a fin de que puedan ser progresivamente más precisos en la calibración de su desempeño.

Palabras clave: juicios metacognitivos; calibración; precisión; nivel de confianza; desempeño; monitoreo metacognitivo.

Resumo

Objetivo: o objetivo deste artigo era realizar uma revisão conceitual da construção "julgamento metacognitivo". **Metodologia**: para a revisão temática da literatura produzida por alguns dos autores mais relevantes na área, 55 fontes documentais foram consideradas e analisadas através das seguintes etapas: introdução, reflexão e estabelecimento de conclusões. **Resultados**: é estabelecido que os julgamentos metacognitivos são entendidos como a construção que agrupa o conjunto de crenças que os estudantes têm sobre o que sabem e não sabem, e também sobre como eles controlam e regulam seu aprendizado. **Conclusão**: o estudo dos julgamentos metacognitivos é apresentado como uma alternativa para favorecer o processo de auto-regulamentação da aprendizagem dos estudantes para que eles possam ser progressivamente mais precisos na calibração de seu desempenho.

Palavras chave: julgamentos metacognitivos; calibração; precisão; nível de confiança; desempenho; monitoramento metacognitivo.

Introduction

A category of research known as "metacognition" emerged as a subject of study in the 1960s, which is when people began to recognize its significance for learning and cognitive performance in areas like problem-solving and critical thinking (Gourgey, 2002; Schraw, 2002; Sawyer, 2014). Since then, research into this construct has been a focus of educational researchers, leading to the current view that it should be included in the curriculum as a cross-cutting subject that should support instruction in all curricular areas (Zohar & Dori, 2012).

In this sense, the student is acknowledged as a self-aware agent who is capable of knowing his or her own thoughts and controlling them. He or she is also capable of assessing their own cognitive performance, motivating themselves, and developing strategies to adjust to changes (Hacker *et al.*, 2009). According to some researchers, students who have greater metacognitive awareness also have greater capacity for developing accurate metacognitive judgments with appropriate levels of self-confidence over their performance (Gutierrez, 2012). As a result, it is believed that the development of metacognitive awareness is essential because it enables students to participate in learning situations while having a greater understanding of and control over their performance. This allows those who create more accurate judgments to develop their monitoring and control skills while also improving their performance and confidence in academic tasks (Gutierrez, 2012).

In this regard, it has been suggested that understanding the elements of metacognitive awareness, which likely serve as the foundation upon which students develop their metacognitive judgments regarding the learning process, may help students reflect on themselves and achieve academic goals (Shaw *et al.*, 2018).

The term "metacognitive awareness" is derived from Flavell's (1979) concept of "metacognition." Researchers have defined it as "the capacity to reflect on one's own learning, understand it, and manage it" (Schraw & Dennison, 1994). "Learning consciousness" refers to the understanding of how to use the information at hand to accomplish a goal, the ability to judge the cognitive demands of a given task, and the assessment of one's progress prior to, during, and after performance (Flavell, 1979; Gourgey, 2002).

Many researchers in the field have reported on Flavell's distinction between knowledge and regulation, which he made in the beginning. They agree that these two are the essential elements of metacognition (Schraw, 2002). However, as presented by Peña-Ayala and Cárdenas (2015), other authors have also posited additional components to those already cited, including skills (Veenman, 2013),

control (Finley *et al.*, 2010), monitoring (Touroutoglou & Efklides, 2010), reflection (Van den Boom *et al.*, 2004), and alertness (Koriat, 2000), among others.

As a result of the pioneering metacognition research, a field of study was established, opening the way for the construct to begin to be connected to other conceptual categories such as affect (Efklides, 2006), cognitive processing (Veenman, 2012), executive control (Schwartz *et al.*, 2013), critical thinking (Ford & Yore, 2012), theory of mind (Flavell, 2004; Misailidi, 2010), cognitive load (Scott & Schwartz, 2007), and motivation (Maier & Richter, 2014).

Research trends on metacognition generally focus on understanding its role in learning processes in various domains of expertise or knowledge, assessing the level of metacognitive awareness among students of various ages, understanding its significance for authorized learning, and, more recently, understanding the role of metacognitive monitoring in relation to student assessments of their confidence in their performance. The purpose of the current article is to provide a general reflection on the calibration issue by describing and analyzing the conceptual underpinnings of the term "metacognitive judgment."

A Current Alternative to Work on Metacognition in Classroom Processes: Metacognitive Judgments

According to Schraw (2009), several terms have been established to address various metacognitive aspects. The study of metacognitive judgments, which is seen as the construct that brings together students' beliefs about what they know and how to control and regulate their learning, is perhaps one of the more novel conceptual categories at the time. This collection of beliefs serves as a legitimate introduction to metacognitive action. In experimental research, judgments about one's own knowledge have been examined in a variety of application areas, including perception, memory, metacognition, decision-making, and work with eyewitnesses (Koriat, 2012).

It is possible to say that specifically in application of the work with metacognitive judgments in education that these are defined as *judgments of probability that inform the student of their own learning and performance before, during, or after certain tests or exams* (Schraw, 2009). According to various authors, the research questions that have received the most attention in the field of studying metacognitive judgments in laboratory and educational settings have been those related to the foundations of judgment, accuracy, reliability, stability, measurement, and control of those judgments (Koriat, 2012; Dunlosky & Thiede, 2013).

In relation to these viewpoints, it has been discovered that students who are diligent in keeping track of their knowledge may discern when they know

something and when they do not. In addition, they can judge when people are rational and when they are confused (Koriat, 2012). The accuracy or precision of the judgments indicates how well a student's judgment corresponds to his actual performance from two important aspects: first, from the degree to which the magnitude of the judgments are related to the actual magnitude of the performance, which is known as "absolute accuracy"; and second, from the degree to which the judgments discriminate between the different levels of performance through the items, that is, the "relative accuracy" (Dunlosky & Thiede, 2013).

In relation to the findings regarding the basis on which students establish the formulation of their metacognitive judgments, according to Koriat (2007), at least three types of perspectives have been pointed out: the direct access approach, the one based on the information, and the experience-based approach.

In Koriat's (2007) opinion, from the direct access approach, metacognitive judgments focus on the activation of an underlying objective. That is, stimuli that are activated in memory increase their strength, which produces high judgments of sensation of knowledge and enhances the precision of decisions. According to the direct access theory approach, people will judge that they know the answer and cannot remember it when the strength of the target is below the recall threshold but above the sense-of-knowledge threshold. If the strength of a target response in memory is below the threshold of awareness sensation, they will judge that they do not recognize the target or stimulus (Dunlosky & Metcalfe, 2009).

Likewise, in relation to the information-based approach, it has been pointed out that in this perspective, emphasis is placed on the content of the specific beliefs and knowledge that the student has about their own skills and competencies. This, to the extent that metacognitive judgments are proposed to be based on the person's theories about how various characteristics of the study material or learning conditions influence memory performance (Koriat, 2007). For example, when students are asked to judge how well they have performed on a test, their judgments may be based on data such as their preconceived notions about their competence in the domain tested, the amount of time they have spent studying for a test or its difficulty, etc. (Koriat *et al.*, 2008).

For its part, the experience-based approach allows for the consideration that mnemonic signals contribute to task performance and, therefore, to constructing judgments while reflecting memories and feelings of knowledge. Trials based on this approach involve a two-stage process. In the first, a subjective feeling is given place; and in the second, that feeling is used to make predictions about memory; for example, by asking the student to evaluate his performance on the test, he may have the experience of detecting the presence of the target, similar to what occurs in the "tip of the tongue" phenomenon, in which the

person may feel that the memory is imminent and may experience frustration at not retrieving the goal that is proving elusive. These feelings can serve as the basis for reported feeling of knowing judgments (*Feeling of Knowing* or *FOK* judgments) (Koriat *et al.*, 2008).

On the other hand, regarding the reliability and stability of metacognitive judgments, authors such as Dunlosky and Thiede (2013) point out that these two aspects have probably been the least addressed at the research level, indicating that, although evidence of stability has been found in judgments in some contexts, generating greater knowledge in this regard is necessary. This, to the extent that students use their confidence judgments to regulate information retrieval, so the use of judgments may have limited effectiveness if they are inaccurate.

In this sense, authors suggest that knowing why stability in the accuracy of the judgments varies is important—for example, between two exams—since this could contribute to the student's achievement. For example, it may be the case that in a first exam some of the students can demonstrate exact accuracy, but in the second exam these same students may present a lower score than in the first (Hadwin & Webster, 2013). This situation reveals the need to continue delving into the link between monitoring and control of decisions in relation to the study (Dunlosky & Thiede, 2013). It can be accomplished with different variables that can influence the accuracy of monitoring—such as the level of confidence and performance—that are related to the person, such as: personality, motivation, and positive or negative emotions; and in relation to the tasks (the difficulty of the item, the length and format of the test, etc.).

Classification of Metacognitive Judgments

One of the aspects related to the understanding and explanation of metacognitive judgments and their importance in learning has to do with their classification or with the approach of the different typologies.

Table 1 presents the classification proposed by Schraw (2009), which is considered the most used classification by researchers in the world. The categories for the typology respond to the moment in which the judgments are made, from a temporal analysis framework and, accordingly, metacognitive judgments can be of the following type: prospective, concurrent, or retrospective (Dunlosky & Metcalfe, 2009; Schraw, 2009).

Table 1. *Taxonomy of Metacognitive Judgments..*

Time of Judgment	Type of Judgment	Description
Prospective (judgments made before the test)	Judgments of learning (JOL).	Judgments about one's ability to retain information.
	Ease of learning judgments (EOL).	Judgments about the relative ease of learning the information.
	Feeling of Knowing judgments (FOK).	Judgments about later recognition of information that could not be recalled.
Concurrent (judgments made during the test)	Online trust judgments[1].	Judgments of confidence in one's own performance.
	Ease of solution judgements.	
	Online judgments of performance accuracy (performance calibration).	Judgments about the accuracy of one's own performance.
Retrospective (judgments made after the test)	Ease of learning/solution.	Post-test judgments about the relative ease of learning information
	Hindsight judgments of performance accuracy, also called "hindsight judgments of confidence"	Judgments about adequate task performance after completing all items.

Source: Schraw (2009, p. 37).

Prospective judgments are predictions that the student makes about performance, which can be evaluated in three ways: a) judgments of learning (JOL), which are predictions about the future performance of the test based on recently studied items or articles (Nelson & Narens, 1994) and imply that the student reviews the information to be learned and makes predictions about how much information will be remembered; b) judgments of feeling of knowledge (FOK), which are presented when the student is asked to predict whether he or she will recognize information that cannot be remembered (these types of judgments assess a person's ability to monitor memory content and ability to retrieve information); and c) judgments of ease of learning (EOL), which involve the ability

1. They are the so-called *learning judgments*.

to judge the amount of time or effort needed to learn the material; that is, they measure the ability to monitor perceived difficulty during the comprehension process (Schraw, 2009).

Alternatively, concurrent judgments are the evaluations made during the learning task. Schraw's (2009) perspective includes confidence judgments, which assess the learner's ability to trust his or her own performance; judgments of ease solution, which refer to the person's ability to monitor the difficulty of the task according to one's cognitive resources; and judgments of accuracy of performance, which measure the subject's ability to monitor his or her performance on the task. In this type of judgments, the student is always asked to give his or her judgment item by item during the test.

In retrospective judgments, the evaluation is made item by item or generally about the set of items on the test after this has been completed. This type of judgment follows the same pattern as the learning judgment, easy to learn/solution, performance, and confidence described previously with the only difference being that the student is asked to make them after taking the test. For this typology, the best known have been called "Judgments of Retrospective Confidence" (JRC).

For the purposes of the present conceptual review, two of the types of judgments are described. The research has been focused on metacognition: learning and confidence judgments (Dunlosky Metcalfe, 2009; Hadwin & Webster, 2013; Narens *et al.*, 2008; Schraw, 2009).

Learning Judgments: a Prospective Type of Judgment

During the last two decades, the approach to learning judgments in classroom processes has been consolidated as an object of research. Developments derived from this type of study have allowed understanding its use in the regulation of study hours, given its function demonstrated in the control of learning that allows people monitoring to guide or indicate which items to study and for how long they need to do it (Dunlosky & Metcalfe, 2009). This typology of judgments allows them to be positioned as a prospective type of judgment in which the person makes predictions about the probability to correctly recover the studied items recently (Dunlosky *et al.*, 2015; Schraw, 2009). This type of judgments has also been called "concurrent," about the execution of the judgments, which can be on-line, while the person performs the task.

The typical experimental format to evaluate JOLs involves asking the person to study the information to be learned (e.g., a list of words), and then

to make predictions about the possibility of recovering each item. This type of judgment involves the person's ability to encode and retain information (Schraw, 2009). The main interest of researchers in the study of learning judgments has been focused on the accuracy of the judgments; that is, the level of the relationship between predicted performance and observed performance on the final test (Narens *et al.*, 2008).

On the accuracy of learning judgments, two factors have received full attention from researchers, as they have found it to be of significant influence: one, the number of trials; and the other, the duration of the trials.

First, the number of trials refers to how extra study sessions seem to improve memory performance (Dunlosky & Metcalfe, 2009). Accordingly, when people study few items and make learning judgments on several tests using the same items, performance on the second test is expected to increase; and with it, the accuracy of the judgments (Koriat, 2000). Thus, at first, one could conclude that people apparently base their learning judgments on the results of the previous trial, which are powerful predictors of performance on the next trial (Finn & Metcalfe, 2014; Vesonder & Voss, 1985).

Second, trial duration as a factor that highly influences learning judgments has been studied as the effect of delayed learning judgments (Dunlosky *et al.*, 2015). This effect consists of asking test takers to make their learning judgments several minutes after having studied the items. The delayed learning trial effect has been replicated several times with high school and university students. Waiting a certain amount of time to make judgments was found to improve judgment accuracy significantly, especially in cases of moderate time delays (Dunlosky & Metcalfe, 2009; Dunlosky *et al.*, 2015).

To understand how learning judgments are developed, several hypotheses have been proposed from the perspective of Dunlosky and Metcalfe (2009). The first of these is the processing facilitation hypothesis, which suggests that the basis of these judgments is the use of heuristics; in other words, the consideration of a type of rule can be valid or incorrect. According to this hypothesis, people make their judgments based on whether the items are easy or not to process, which would lead them to make more accurate judgments about performance.

The second hypothesis is the memory fluency hypothesis, which also assumes that learning judgments are heuristic by nature and that memory answer is a sign of good memory capacity. In this way, it is considered that when information is kept in mind more quickly, then memory is fluid. Finally, the hypothesis of the use of cues is recognized, from which it is proposed that people's learning judgments are significantly different according to three types of cues. In Koriat's (2000) perspective, these cues can be intrinsic (as characteristics of the items that may decrease the learning difficulty), extrinsic (those involving

encoding ability or other non-intrinsic learning conditions for the items), and mnemonic (understood as internal cues based on subjective experiences that suggest that an item will be remembered).

Learning Judgments: a Concurrent Type of Judgment

Dunlosky and Metcalfe (2009) pose that Judgments of Confidence (JOCs) require people to rate the probability that their answers are correct. The confidence people have in their beliefs and knowledge is considered important. This defines whether an answer is shared or not, considering whether others believe in it. According to Hadwin and Webster (2013), JOCs are considered indicators of metacognitive monitoring because they represent students' perception or awareness of their own cognitive processes.

According to Winne (2011), confidence judgments exert self-regulatory action by triggering If-Then-Else sequences, implying that if the learner does not feel safe, then they will adjust their expectations or do something to improve their chances of being successful on the task. In the same way, if one does not reach the goal, one may be less confident in future judgments or may set goals that are easier to achieve. These aspects represent the importance of confidence judgments for control and metacognitive regulation.

In the same approach, Dinsmore and Parkinson (2013) note that students' confidence ratings include both person elements (such as their prior knowledge) and task elements (e.g., text features), which led them to conclude that students base their confidence on a combination of person and environmental characteristics. These results are consistent with findings derived from studies in self-efficacy, in which confidence judgments emerge in response to past experiences and in which personal performance and achievements are a prominent source for constructing new judgments (Hadwin & Webster, 2013).

Like the basis that explains how students construct their metacognitive judgments, Koriat (2012) posits that feelings of confidence may be based on recognition of the question (direct access approach), inferences about the task at hand (information-based approach), memories and feelings of knowledge derived from the task (experience-based approach), or memories of past performance rather than content-specific knowledge (self-consistency approach).

The confidence judgments raised by students are rarely considered to be perfect, which triggers two biases, known as "overconfidence" and

"under-confidence" (Dunlosky & Metcalfe, 2009). In this regard, Hadwin and Webster (2013) consider that overconfidence may lead students to fail to recognize when to regulate strategies to increase their successful performance, while under-confidence would lead to the unnecessary use of cognitive and affective resources to achieve goals that have already been attained.

Explanatory Models of Metacognitive Judgments that Highlight the Importance of Metacognitive Monitoring

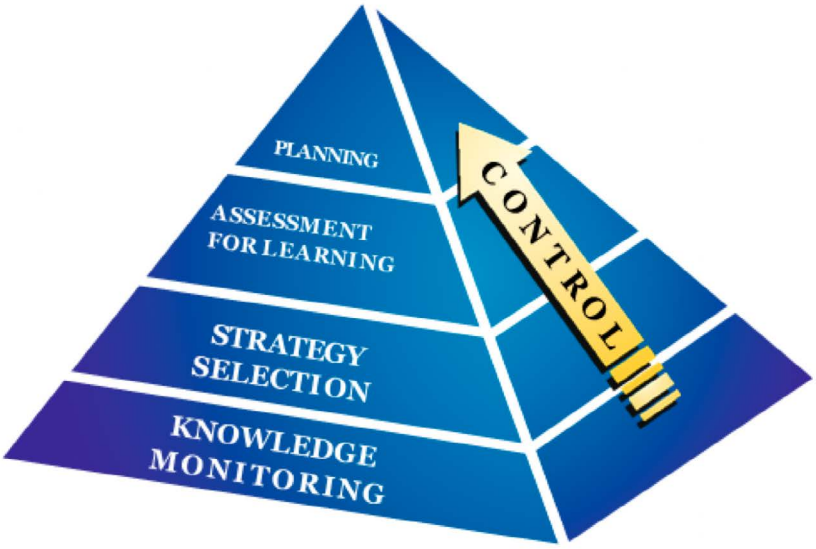
Tobias and Everson (2002) Model.

This model emphasizes monitoring before learning as a requirement for the metacognitive process, which implies the ability to evaluate learning, choose strategies, and make plans for one's learning process, as shown in Figure 1.

In the perspective of Tobias and Everson (2002), three components are recognized as necessary to regulate learning effectively: knowledge of metacognition, the ability to monitor learning processes, and the meta-ability to control such processes.

The basic premise of Tobias and Everson's model (2002) is that knowledge monitoring is the ability to know what one knows, as well as knowing what one does not know. Thus, students who correctly distinguish between what they have learned and what they still need to learn have an advantage during instruction because they can skip the more familiar material and focus on the less familiar content they have yet to master; whereas students with less metacognitive awareness, i.e., those with less accurate knowledge monitoring skills, often spend too much time reviewing familiar material at the expense of mastering unfamiliar or new content (Tobias & Everson, 1996; 2009).

Figure 1. *Hierarchy of Metacognitive Processes.*



Source: Tobias & Everson (2009, p. 115).

Nelson and Narens (1990) Model.

According to Schraw *et al.* (2013), in studies on metacognition the model of Nelson and Narens (1990) has served as an initial theoretical framework for its conceptualization. Hacker *et al.* (2009) argue that three aspects underlie their model:

a) Mental processes are divided into two or more specifically interrelated levels: a cognitive level and a metacognitive level.

b) The metacognitive level contains a representation or dynamic mental model of the cognitive level; and in turn, the cognitive level is responsible for the activity of the cognitive processes themselves.

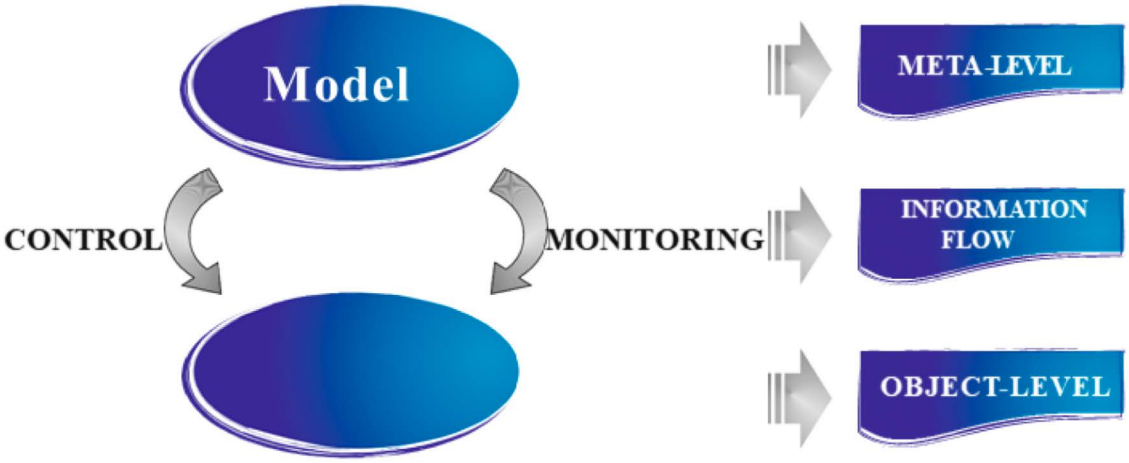
c) There are two reciprocal relationships between the two levels: monitoring and control. These are defined in terms of the direction of information flow between the meta-level and the object-level (p. 161).

The elements that form the model include two levels of metacognition. One is the object-level, characterized by the mastery of task-relevant knowledge and the use of a repertoire of automated strategies. The other is the meta-level, characterized by an explicit mental model of strategy use that controls and regulates learning (Schraw & Gutierrez, 2015).

Monitoring refers to the type and quality of information received from the object-level so that the target level can make the necessary changes. That is, monitoring is the process through which the learner uses information from the object-level to assess progress toward a learning goal at the meta-level (Gutierrez, 2012). This process involves information gathering and represents the ongoing flow between the meta-level and the object-level, allowing the learner to construct plans and assess meta-level accuracy or performance when carrying out a learning task.

In turn, the control, which can also be understood as an executive process, involves the interventions that students make in their environment to achieve a goal and indicates the ability of the meta-level to make adaptations at the object-level. Thus, the meta-level reacts to stimuli by generating a control flow that starts, alters, or finishes the mental actions performed at the object-level (Nelson & Narens, 1994; Hacker *et al.*, 2009). Figure 2 shows the representation of the Nelson and Narens (1994) model.

Figure 2. *Representation of the Nelson and Narens Metacognition Model (1990).*



Source: Nelson & Narens (1990); Gutierrez (2012, p. 7).

To exemplify this model, Dunlosky and Thiede (2013) posit that when students are studying for an exam, they may decide to start by reading the assigned chapters. As they do so, they can evaluate their progress and judge how well they understand the concepts, so that if they judge they have learned them well, they can stop when they feel it is necessary. Conversely, if they judge that they do not understand some concepts, they will resort to strategies such as rereading chapters or seeking help from peers or teachers. According to Dunlosky and Thiede's (2013) perspective, these monitoring and control mechanisms may influence students' achievements, on the one hand, because those who are overconfident in their understanding may not be sufficiently prepared, as they prematurely culminate their study hours; and on the other hand, because, even if students' judgments are accurate, if they do not use them effectively to control their learning they may even obtain low results.

From this perspective, Gutierrez (2012) notes that thinkers with greater metacognitive awareness have more refined monitoring (information-gathering) of the environment (the object-level) that continuously informs the meta-level model (environment representation) about the state of the object-level. Thus, these students have a more perfect representation of the environment and a greater awareness of their metacognitive processes, so they will have a greater ability to produce more accurate metacognitive judgments with more appropriate levels of confidence regarding their performance than their less metacognitively aware peers (Gutierrez, 2012).

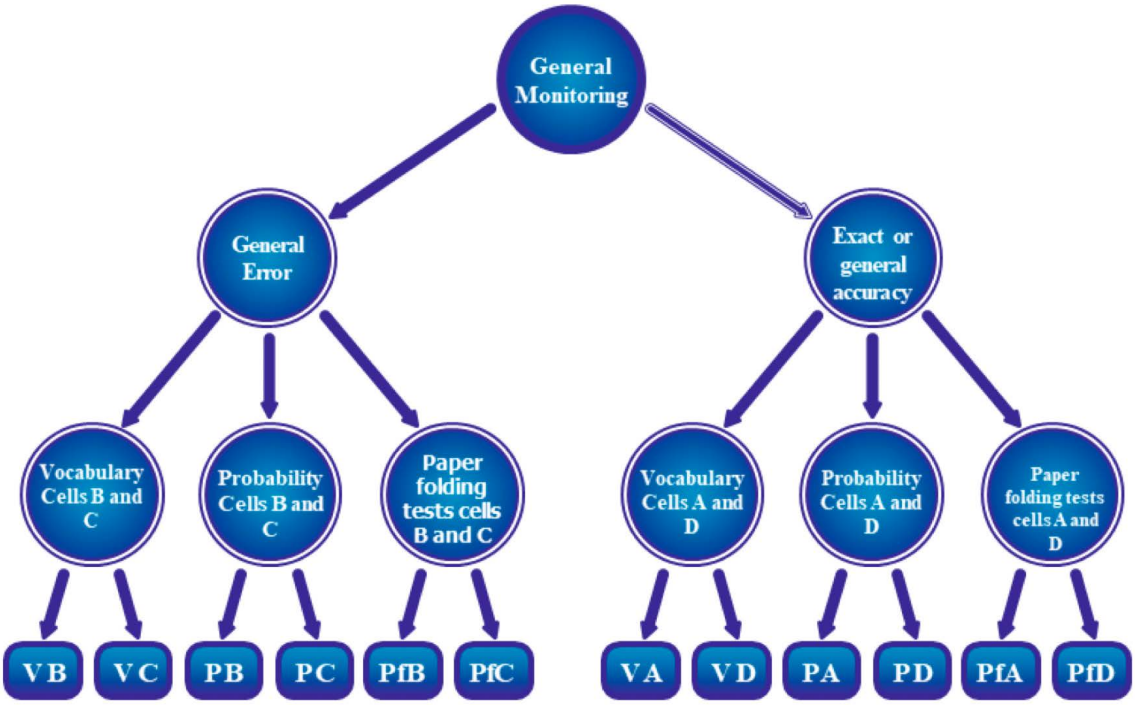
Third-order General Monitoring Model (Schraw et al., 2013; Gutierrez *et al.*, 2016; 2021)

This model demonstrates the existence of two different types of metacognitive monitoring: one for accuracy and the other for errors. Thus, metacognitive judgments derived from accurate monitoring are different from those formulated from error; a form of processing that, in turn, can be evidenced in overconfidence or under-confidence biases (Gutierrez *et al.*, 2016).

Gutierrez *et al.* (2016), through factor analysis, provided evidence for the existence of two different factors involved in metacognitive monitoring: overall accuracy and overall error (see Figure 3). Thus, they found error and accuracy factors specific to certain domains (vocabulary, paper folding, and probability tasks) (first order), which loaded to domain-general error and accuracy factors (second order), which then loaded to a general monitoring factor (third order). Considering that from this theoretical model it has been shown that overall

accuracy and error are treated as separate latent dimensions, future interventions in metacognitive monitoring should be oriented toward improving accuracy or decreasing error, but not in the direction of both goals necessarily (Gutierrez, 2020; Gutierrez *et al.*, 2020).

Figure 3. *Representation of the Gutierrez et al. (2016) General Third-order Monitoring Model.*



Source: Gutierrez *et al.* (2016, p. 4).

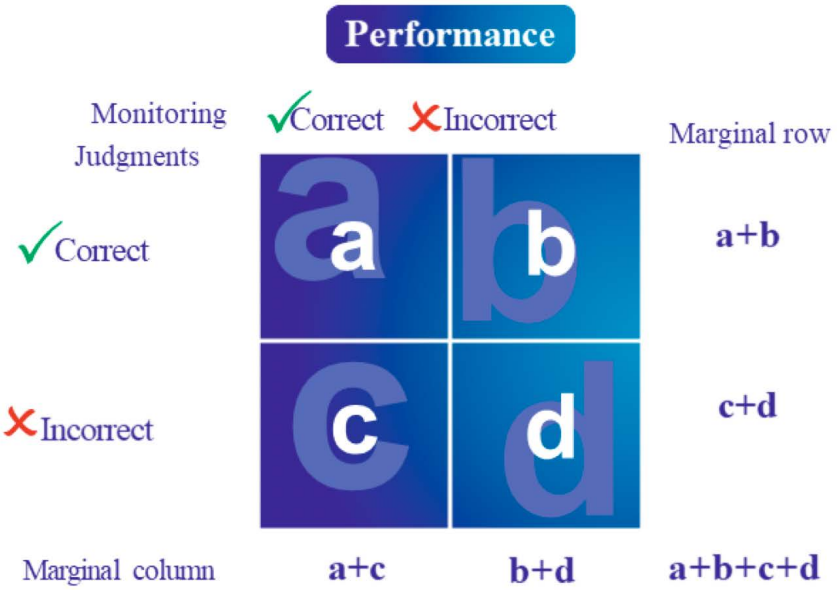
The Role of Metacognitive Judgments in Metacognitive Monitoring.

Schwartz and Perfect (2004) state that metacognitive monitoring refers to the processes that allow individuals to observe, experience, or reflect on their own cognitive processes. These processes also aid in knowing how much or how little to be learned, whether individuals have mastered, for example, the multiplication tables, or the feeling of having understood the text they have just read. Therefore, it has been indicated that monitoring is the learners' ability to successfully judge their own cognitive processes as well as their own performance.

Additionally, metacognitive monitoring has been understood as the relationship between task performance and judgment about that performance (Gutierrez *et al.*, 2016). Such relationship has been researched in studies on calibration that is the process of accuracy or alignment between a judgment and a meaningful standard which is the performance on a given task or performance test (Dinsmore & Parkinson, 2013; Hadwin & Webster, 2013). Metacognitive monitoring is exhibited by asking students to make metacognitive judgments about their assessments or assignments, and therefore, they may know about the state of their cognition.

The traditional procedure for studying metacognitive monitoring involves completing a test item and asking the students to judge their performance on answers. According to Gutierrez *et al.* (2016), in their research on a 2 x 2 matrix, the four monitoring outcomes that arise after the students complete the test items and make the metacognitive judgments is depicted in Figure 4.

Figure 4. *Relationships between the Four Types of Performance-Judgements Outcomes.*



Source: Gutierrez *et al.* (2016, p. 2).

In line with Gutierrez *et al.* (2016), the double-entry matrix is proposed for the analysis of the two monitoring mechanisms: hits and misses processing. Cell *a* shows the first outcome and corresponds to correct performance that is judged as "correct," i.e. accurate monitoring. Cell *d* represents incorrect performance that is judged as incorrect, which would also imply an accurate monitoring of the process itself. Cell *b* indicates an incorrect performance that is judged as correct; and cell *c* reflects a correct performance that is judged as incorrect, representing inaccurate monitoring. This inaccurate monitoring has been characterized as "over or under-confidence," or as "illusion of knowing" and "illusion of not knowing" (Gutierrez *et al.*, 2016; Gutierrez *et al.*, 2020; Gutierrez, 2020).

According to the findings from the third-order general monitoring model, metacognitive monitoring can be explained by either the domain-specific monitoring hypothesis or the domain-general monitoring hypothesis. The first hypothesis suggests that monitoring accuracy is situated within a specific content domain, for example, mathematics as the domain, algebra as the subdomain or a task, such as proofreading. Similarly, the domain-general monitoring hypothesis

posits that students construct a repertoire of general skills that allow them to make accurate judgments about their performance. This repertoire may comprise skills such as goal setting, strategy management, and self-explanation, among others (Gutierrez *et al.*, 2016; Gutierrez *et al.*, 2020).

Evaluation of Metacognitive Judgments.

Judgment accuracy is one of the most studied metacognitive judgments, that is, with the question "how well are the people's judgments related to their real performance?" (Dunlosky & Thiede, 2013).

In this regard, Hadwin and Webster (2013) state that metacognitive judgments have been studied as predictions made prior to completing a task or question. For example, by asking the individuals to indicate their degree of confidence in their ability to recall the second word in tasks of associated pairs (Nelson & Dunlosky, 1991); and, as predictions after completing the task. This time by asking how confident they are that their answers are correct (Dinsmore & Parkinson, 2013). Similarly, as global predictions, in which the percentage of questions answered correctly have been examined, and as local predictions, in which confidence for a specific item or question is examined (Pieschl, 2009).

In this regard, Schraw (2009) states that most of the studies on metacognitive monitoring focus on the relationship between metacognitive judgments and performance. For this purpose, measures of absolute accuracy and relative accuracy have been used.

Absolute accuracy measures whether a metacognitive judgment exactly matches performance, providing a measure of how accurately a person can judge performance on the test item. This implies that the students' confidence judgments are compared in an absolute fashion to their performance on the same task (Gutierrez, 2012). Relative accuracy, on the other hand, provides a measure of both the relationship between correct and incorrect metacognitive judgments and a set of metacognitive judgments and the results on a performance test. This allows describing the consistency of judgments or how well individuals can discriminate better learned material from less learned material (Schraw, 2009). This type of accuracy is assessed using correlational coefficient measures, such as "Pearson" or "Gamma" (Nelson & Narens, 1994).

Table 2 depicts the traditional classification of the measures in the statistical analysis of judgments and the interpretation for each score.

Table 2. . *Types of Outcome Measures.*

Measured Construct	Outcome Measure	Interpretation of Punctuation.
Absolute accuracy	Absolute accuracy index	Discrepancy between a confidence judgment and performance.
	Hamann Coefficient	Discrepancy between hits and misses in a contingency table.
	Bias Index	The degree of over or under- confidence in judgments.
Relative accuracy	Correlation Coefficient	Relationship between set of judgments and corresponding performances scores.
	Gamma Coefficient	Dependence between judgments and performance.
	Discrimination Index	Ability to discriminate between correct and incorrect outcomes.

Source: Schraw (2009, p. 39).

These outcome scores are calculated based on different computational formulas and statistical measures that combine the information of the four cells for an estimation of the calibration process (Schraw *et al.*, 2013). Among the most commonly used measures are: G-index, Gamma, d', and sensitivity and specificity. However, as Schraw *et al.* (2014) state, researchers tend to consider them as mutually exclusive. This is framed within a current debate. From the perspective of Nelson (1984), the Gamma measure has been proposed as superior to the other measures, indicating the problem as unidimensional; while from Schraw's (1995) initial findings and many of his later works (Schraw *et al.*, 2013; Schraw *et al.*, 2020), the problem of measuring judgments is regarded as multidimensional, therefore the need to combine the power of different statistical measures.

Conclusions

Calibration studies address the relationship between metacognition and the students' own performance as an alternative to favor self-regulated learning. Therefore, calibration is understood as the degree to which individuals judge their performance and how this judgment or belief corresponds or adjusts to their actual level of performance (Gutierrez, 2012).

Studies on calibration and metacognition to achieve and/or improve the students agency capacity and self-regulation of learning have been considered important because calibration is assumed to measure important attributes of effective metacognitive monitoring (Nelson, 1996), among which confidence and accuracy of metacognitive judgments could be considered. Traditionally, studies in this line have been oriented to review calibration processes in text comprehension, multimedia learning, test preparation, learning in computer environments, as well as in collaborative learning processes (Hacker *et al.*, 2008; Hacker *et al.*, 2009; Zimmerman & Moylan, 2009; Winne & Azevedo, 2014).

This article aimed to provide a comprehensive view of the problem of calibration from the description and analysis of some of the conceptual aspects of the construct "metacognitive judgment," understood as an inductor of metacognitive action and to favor the students' calibration accuracy with respect to their own learning process.

References

Dinsmore, D., L. & Parkinson, M., M. (2013). What are Confidence Judgments Made of? Students' Explanations for their Confidence Ratings and what that Means for Calibration. *Learning and Instruction*, *24*, 4-14. https://doi.org/10.1016/j.learninstruc.2012.06.001

Dunlosky, J. & Thiede, K., W. (2013). Four Cornerstones of Calibration Research: Why Understanding Students' Judgments can Improve their Achievement. *Learning and Instruction*, *24*, 58-61. https://doi.org/10.1016/j.learninstruc.2012.05.002

Dunlosky, J & Metcalfe, J. (2009). *Metacognition*. SAGE Publications Ltd.

Dunlosky, J., Serra, M., J., Matvey, G. & Rawson, K., A. (2015). Second-Order Judgments About Judgments of Learning. *The Journal of General Psychology*, *132*(4), 335–346. https://doi.org/10.3200/GENP.132.4.335-346

Efklides, A. (2006). Metacognition and Affect: What can Metacognitive Experiences Tell Us About the Learning Process? *Educational Research Review*, *1*, 3-14. https://doi.org/10.1016/j.edurev.2005.11.001

Flavell, J., H. (1979). Metacognition and Cognitive Monitoring: A New Area of Cognitive–Developmental Inquiry. *American Psychologist*, *34*(10), 906-911. https://doi.org/10.1037/0003-066X.34.10.906

Flavell, J. (2004). Theory of Mind Development: Retrospect and Prospect. *Merrill-Palmer Quarterly*, *50*(3), 274-290. https://doi.org/10.1353/mpq.2004.0018

Finley, J. R., Tullis, J. G., & Benjamin, A. S. (2010). Metacognitive Control of Learning and Remembering. In *New Science of Learning: Cognition, Computers and Collaboration in Education* (pp. 109-131). Springer. https://doi.org/10.1007/978-1-4419-5716-0_6

Finn, B. & Metcalfe, J. (2014). Overconfidence in Children's Multi-Trial Judgments of Learning. *Learning and Instruction*, *32*, 1-9. https://doi.org/10.1016/j.learninstruc.2014.01.001

Ford, C. L. & Yore, L. D. (2012). Toward Convergence of Critical Thinking, Metacognition and Reflection: Illustrations from Natural and Social Sciences, Teacher Education, and Classroom Practice. En A., Zohar & Y.J., Dori, (Eds.), *Metacognition in Science Education: Trends in Current Research, Contemporary Trends and Issues in Science Education*, (pp. 251-272). Springer. http://dx.doi.org/10.1007/s10763-016-9726-x.

Gourgey, A. (2002). Metacognition in Basic Skills Instruction. In H. Hartman (Ed.), *Metacognition in learning and instruction* (pp. 17-32). Springer–Science. https://doi.org/10.1007/978-94-017-2243-8

Gutierrez, A. (2012). *Enhancing the Calibration Accuracy of Adult Learners: A Multifaceted Intervention* (tesis doctoral). University of Nevada, Las Vegas. http://digitalscholarship.unlv.edu/cgi/viewcontent.cgi?article=2569&context=thesesdissertations

Gutierrez, A., Schraw, G., Kuch, F. & Richmond, A. (2016). A Two-Process Model of Metacognitive Monitoring: Evidence for General Accuracy and Error Factors. *Learning and Instruction*, *44*, 1-10. https://doi.org/10.1016/j.learninstruc.2016.02.006

Gutierrez, A., Schraw, G., Kuch, F. & Richmond, A. (2021). General Accuracy and General Error Factors in Metacognitive Monitoring and the Role of Time on Task in Predicting Metacognitive Judgments. *Revista de Psicología CES*, *21*(1), 1-42. https://www.researchgate.net/publication/346408081_General_Accuracy_and_General_Error_Factors_in_Metacognitive_Monitoring_and_the_Role_of_Time_on_Task_in_Predicting_Metacognitive_Judgments

Gutierrez de Blume, A. P. (2020). Efecto de la instrucción de estrategias cognitivas en la precisión del monitoreo metacognitivo de los alumnos universitarios estadounidenses. *Revista Tesis Psicológica*, *15*(2), 1-29. https://doi.org/10.37511/tesis.v15n2a12

Hacker, D., Bol, L. & Bahbahani, K. (2008). Explaining Calibration Accuracy in Classroom Contexts: The Effects of Incentives, Reflection, and Explanatory Style. *Metacognition Learning*, *3*, 101–121. https://doi.org/10.1007/s11409-008-9021-5

Hacker, D., Dunlosky, J. & Graesser, A., C. (2009). A Growing Sense of "Agency". In D. Hacker, J. Dunlosky y A. Graesser (eds.), *Handbookof metacognition education* (pp. 1-4). Routledge Taylor & Francis Group.

Hacker, D., J., Keener, M., C. & Kircher, J., C. (2009). Writing is Applied Metacognition. In D. Hacker, J. Dunlosky & A. Graesser (Eds.), *Handbook of Metacognition in Education* (pp. 154-172). Routledge Taylor & Francis Group.

Hadwin, A., F. & Webster, E., A. (2013). Calibration in Goal Setting: Examining the Nature of Judgments of Confidence. *Learning and instruction*, *24*, 37-47. https://doi.org/10.1016/j.learninstruc.2012.10.001

Koriat, A. (2000). The Feeling of Knowing: Some Metatheoretical Implications for Consciousness and Control. *Consciousness and Cognition*, *9*, 149-171. https://doi.org/10.1006/ccog.2000.0433

Koriat, A. (2007). Metacognition and Consciousness. In P. D. Zelazo, M. Moscovitch, & E. Thompson (Eds.), *The Cambridge Handbook of Consciousness* (pp. 289-325). Cambridge University Press. https://doi.org/10.1017/CBO9780511816789.012

Koriat, A., Nussinson, R., Bless, H. & Shaked, N. (2008). Information-Based and Experience- Based Metacognitive Judgments: Evidence from Subjective Confidence. En J., Dunlosky & R., Bjork (Eds.). *Handbook of Metamemory and Memory* (pp. 117-136). Psychology Press.

Koriat, A. (2012). The Subjective Confidence in One's Knowledge and Judgements: Some Metatheoretical Considerations. En M., J., Beran, J., L. Brandl, J., Perner & J., Proust (Eds), *Foundations of Metacognition* (pp. 504-555). Oxford University Press.

Maier, J. & Richter, T. (2014). Fostering Multiple Text Comprehension: How Metacognitive Strategies and Motivation Moderate the Text-Belief Consistency Effect. *Metacognition and Learning*, *9*(1), 51–74. https://doi.org/10.1007/s11409-013-9111-x

Misailidi, P. (2010). Children's Metacognition and Theory of Mind: Bridging the Gap. In A., Efklides & P., Misailidi, (Eds.). *Trends and Prospects in Metacognition Research* (pp. 279–291). Springer Science+Business Media. https://doi.org/10.1007/978-1-4419-6546-2_13

Narens, L., Nelson, T. O., & Scheck, P. (2008). Memory Monitoring and the Delayed JOL Effect. In J. Dunlosky & R. A. Bjork (Eds.), *Handbook of Metamemory and Memory* (pp. 137-153). Psychology Press.

Nelson, T. (1996). Gamma is a Measure of the Accuracy of Predicting Performance on One Item Relative to Another Item, Not of the Absolute Performance on an Individual Item. *Applied Cognitive Psychology*, *10*, 257–260. https://doi.org/10.1002/(SICI)1099-0720(199606)10:3<257::AID-ACP400>3.0.CO;2-9

Nelson, T,O., & Narens, L. (1990). Metamemory: A Theoretical Framework and New Findings. *The Psychology of Learning and Motivation*, *26*, 125-173. https://doi.org/10.1016/S0079-7421(08)60053-5

Nelson, T. O., & Dunlosky, J. (1991). When People's Judgments of Learning (JOLs) are Extremely Accurate at Predicting Subsequent Recall: The "Delayed-JOL Effect". *Psychological Science*, *2*, 267-270. https://doi.org/10.1111/j.1467-9280.1991.tb00147.x

Nelson, T.O. & Narens, L. (1994). Why Investigate Metacognition? In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition* (pp. 1–25). The MIT Press.

Peña-Ayala, A. & Cárdenas, L. (2015). A Conceptual Model of Metacognitive Activity. En A., Peña-Ayala (Ed.), *Metacognition: Fundaments, Applications, and Trends: A profile of the current state of the art* (pp. 39-64). Springer.

Pieschl, S. (2009). Metacognitive Calibration and Extended Conceptualization and Potential Applications. *Metacognition and Learning*, *4*, 3-31. https://doi.org/10.1007/s11409-008-9030-4.

Sawyer, R., K. (2014). Introduction: The New Science of Learning. In R., K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 1-20). Cambridge University Press.

Schraw, G. & Dennison, R. S. (1994). Assessing Metacognitive Awareness. *Contemporary Educational Psychology*, *19*, 460-475. https://doi.org/10.1006/ceps.1994.1033

Schraw, G. (2002). Promoting General Metacognitive Awareness. In H. Hartman (Ed.), *Metacognition in Learning and Instruction* (pp. 3- 16). Springer – Science.

Schraw, G. (2009). Measuring Metacognitive Judgment. In D. Hacker, J. Dunlosky & A. Graesser (Eds.), *Handbook of Metacognition in Education* (pp. 415-429). Routledge Taylor & Francis Group.

Schraw, G., Kuch, F. & Gutierrez, A.P. (2013). Measure for Measure: Calibrating Ten Commonly Used Calibration Scores. *Learning and Instruction*, *24*(48-57). https://doi.org/10.1016/j.learninstruc.2012.08.007

Schraw, G., Kuch, F., Gutierrez, A.P. & Richmond (2014). Exploring a Three-Level Model of Calibration Accuracy. *Journal of Educational Psychology*, *106*(4), 1192–1202. https://doi.org/10.1037/a0036653.

Schraw, A. & Gutierrez, A. (2015). Metacognitive Strategy Instruction that Highlights the Role of Monitoring and Control Processes. In: A. Peña-Ayala (Ed.), *Metacognition: Fundaments, Applications and Trends* (pp. 3-15).: Springer International Publishing Switzerland.

Schwartz, N. H. & Perfect, T. (2004). Introduction: Toward an Applied Metacognition. En T., J. Perfect & D., S. Lindsay (Eds.), *The SAGE Handbook of Applied Metacognition* (pp. 3-19). SAGE Publications Inc.

Schwartz, N.H., Scott, B.M. & Holzberger, D. (2013). Metacognition: A Closed-Loop Model of Biased Competition–Evidence from Neuroscience, Cognition, and Instructional Research. En R., Azevedo & V., Aleven, (Eds.) *International Handbook of Metacognition and Learning Technologies* (pp. 79-94). Springer International Handbooks of Education. https://doi.org/10.1007/978-1-4419-5546-3_6

Scott, B.M. & Schwartz, N.H. (2007). Navigational spatial displays: the role of metacognition as cognitive load. L*earning and Instruction, 1*(17), 89–105. https://doi.org/10.1016/j.learninstruc.2006.11.008

Shaw, S., Kuvalja, M., & Suto, I. (2018). An Exploration of Nature and Evaluation of Student Reflection. *Research Matters: A Cambridge Assessment publication, 25, 2-8.*

Tobias, S. & Everson, H. T. (1996). *Assessing Metacognitive Knowledge Monitoring.* College Entrance Examination Board.

Tobias, S. & Everson, H. T. (2002). The Ability to Estimate Knowledge and Performance in College: A Metacognitive Analysis. En H. J. Hartman (Ed.), *Metacognition in learning and instruction: Theory, research and practice* (pp. 69-84). Kluwer Academic. https://doi.org/10.1007/978-94-017-2243-8

Tobias, S. & Everson, H., T. (2009). The Importance of Knowing What You Know: A Knowledge Monitoring Framework for Studying Metacognition in Education. In D. Hacker, J. Dunlosky & A. Graesser (Eds.), *Handbook of Metacognition in Education* (pp. 107-128). Routledge Taylor & Francis Group.

Touroutoglou, A. & Efklides, A. (2010). Cognitive Interruption as an Object of Metacognitive Monitoring: Feeling of Difficulty and Surprise. En

A., Efklides & P., Misailidi, (Eds.). *Trends and Prospects in Metacognition Research* (pp. 171-208). Springer Science+Business Media.

Van den Boom, G., Paas, F., Van Merriënboer, J. J. G., & Van Gog, T. (2004). Reflection Prompts and Tutor Feedback in a Web-Based Learning Environment: Effects on Students Self-Regulated Learning Competence. *Computers in Human. Behavior*, *20*, 551–567. https://doi.org/10.1016/j.chb.2003.10.001

Veenman, M., J. (2012). Metacognition in Science Education: Definitions, Constituents, and Their Intricate Relation with Cognition. In A., Zohar & Y.J., Dori, (Eds.), *Metacognition in Science Education: Trends in Current Research*, *Contemporary Trends and Issues in Science Education* (pp. 21-36). Springer.

Veenman, M.V.J. (2013). Assessing Metacognitive Skills in Computerized Learning Environments. En R., Azevedo & V., Aleven, (Eds.) *International Handbook of Metacognition and Learning Technologies* (pp. 157-168). Springer Science+Business Media.

Vesonder, G. T., & Voss, J. F. (1985). On the Ability to Predict One's Own Responses While Learning. *Journal of Memory and Language*, *24*, 363–376. https://doi.org/10.1016/0749-596X(85)90034-8

Winne, P. H. (2011). A Cognitive and Metacognitive Analysis of Self-Regulated Learning. En B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 15-32). Routledge.

Winne, P. H, & Azevedo, R. (2014). Metacognition. In: T. Sawyer. (Ed), *The Handbook of the Learning Sciences* (pp. 63- 87). Cambridge University Press.

Zohar, A. & Dori, Y., J. (2012). Metacognition in Science Education Trends in Current Research. In A. Zohar & Y. J. Dori (Eds.), *Metacognition in Science Education: Trends in Current Research*, *Contemporary Trends and Issues in Science Education* (pp. 1-20). Springer.

Zimmerman, B. J. & Moylan, A. R. (2009). Self-Regulation: Where Metacognition and Motivation Intersect. In: D. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Handbook of Metacognition in Education* (pp. 299-315). Routledge.