Differences in Metacognitive Skills among Undergraduate Students in Education, Psychology, and Medicine

ANTONIO P. GUTIÉRREZ DE BLUME

Georgia Southern University, Georgia, United States of America

DIANA MARCELA MONTOYA LONDOÑO

Universidad de Caldas y Universidad de Manizales, Manizales, Colombia

Abstract

Metacognitive skills such as when and why to apply strategies successfully given task demands (conditional knowledge) and those that assist in regulation like comprehension monitoring are essential for effective learning. However, the debate regarding whether metacognitive skills are domain general or domain specific continues to rage among scholars. Presumably, if metacognitive skills are domain specific, there should be significant differences between domains whereas if they are domain general, there should be no differences across domains. Thus, in the present study we examined the generality/specificity of metacognitive skills (knowledge of cognition: declarative, procedural, and conditional; regulation of cognition: planning, information management, debugging, comprehension monitoring, and evaluation) in a sample of Colombian university students ($N = 507$) studying education ($n = 156$), psychology ($n = 166$), and medicine ($n = 185$) employing the Spanish version of the Metacognitive Awareness Inventory. Results revealed that there were significant differences in all but two metacognitive skills (procedural knowledge and debugging) across domains, largely supporting the domain specific hypothesis, but also partially supporting the domain general view. Implications and recommendation of the findings for theory, research, and practice are discussed.

Keywords: domain general, domain specific, metacognitive knowledge, metacognitive regulation, metacognition, self-regulation.

Diferencias en Habilidades Metacognitivas entre Estudiantes de Pregrado en Educación, Psicología y Medicina

Resumen

Entre las habilidades metacognitivas se encuentran la capacidad para saber cómo, cuándo y por qué aplicar estrategias de forma exitosa según las demandas de la tarea (conocimiento condicional) y aquellas otras capacidades que ayudan en la regulación cognitiva, como la supervisión en línea de la actividad de aprendizaje en curso, y la comprensión (monitoreo). Dichas habilidades se consideran esenciales para un aprendizaje efectivo. Sin embargo, el debate sobre si las habilidades metacognitivas son de dominio general o específico continúa siendo un tema de controversia entre los académicos. Presumiblemente, si estas habilidades son específicas de un dominio, debería haber diferencias significativas entre dominios, mientras que, si son de dominio general, estas diferencias no deberían de estar presentes. De acuerdo con lo anterior, en el presente estudio se examinó la generalidad/especificidad de las habilidades metacognitivas (conocimiento de la cognición: declarativo, procedimental y condicional; regulación de la cognición: planificación, gestión de la información, monitoreo, depuración y evaluación) en una muestra colombiana de estudiantes universitarios ($N = 507$) que, para el momento del estudio, se encontraban cursando un programa de pregrado en educación ($n = 156$), psicología ($n = 166$) y medicina ($n = 185$), mediante el uso de la versión en español del *Inventario de Conocimiento Metacognitivo*. Los resultados revelaron diferencias significativas en todas las habilidades metacognitivas evaluadas para las diferentes áreas de dominio, con excepción de dos habilidades metacognitivas: conocimiento procedimental y depuración, lo que respalda en gran medida la hipótesis específica del dominio, pero también apoya parcialmente la vista general del dominio. Se discuten las implicaciones y recomendaciones de los hallazgos para la teoría, la investigación y la práctica.

Palabras clave: autorregulación, conocimiento metacognitivo, dominio general, dominio específico, metacognición, regulación metacognitiva.

IN THE context of metacognition research, at least three historical perspectives are recognized. They are distinguished in the theoretical position of different researchers in the field, in relation to the approach to the discussion about whether metacognition corresponds to a set of general metacognitive skills or if these are developed as a function of domain of the learning task. Thus, an initial perspective is recognized from where metacognition is postulated as a set of skills of general mastery, more at the level of a *metacapacity* for the agency and regulation of one's own learning, that generally serves to understand and regulate one's own metacognitive activity, regardless of the domain under investigation. In this perspective is the classical theory about general metacognitive awareness (Schraw, 2002) that posits the existence of a general metacognitive capacity that allows individuals to know and regulate their own learning process, regardless of the domain.

Historical Approaches to the Study of Metacognition

This section focuses on three historical approaches to research on metacognition. The first approach endeavors to explain the nuances of the micro-processes involved in metacognition while the second one attempts to understand how domain generality and specificity influence these micro-processes. Finally, the third approach employs more sophisticated statistical modeling techniques for a better measurement of various aspects of metacognition.

Classical conceptual approaches to metacognition are understood as the capacity that allows students to know themselves and consciously inspect their own cognition and its attributes such as how they think and with what information they think, then charting and then mentally following a path of action to achieve a learning goal (Winne & Marzouk, 2019). In this sense, Schraw (2002) has argued that cognition and metacognition differ in that cognitive skills are necessary to perform a task while metacognition is necessary

to understand how the task was executed (Garner, 1987; Schraw, 2002). Similarly, Schraw's (2002) theory about general metacognitive awareness states that knowledge and metacognitive regulation are recognized as components of a multidimensional general domain that are teachable and malleable. Thus, metacognitive knowledge and regulation span a wide variety of subject areas and domains, with empirical evidence to support the conclusion that students demonstrate a general monitoring skill that evolves from tacit to informal, to formal actions of metacognition (Schraw & Moshman, 1995). Schraw (2002) postulated that cognitive skills were encapsulated within domains or subject areas, while metacognitive skills covered multiple domains, multiple tasks and exhibited flexibility in new learning tasks, including in domains that may have little in common (Schraw & Moshman, 1995).

Schraw (2002) initially considered that metacognitive knowledge was in principle domain or task specific and that as students acquired more metacognitive knowledge in various domains, they might be more capable of constructing general metacognitive knowledge such as understanding memory limitations (i.e., metamemory) and also general regulatory skills, like selecting appropriate learning strategies (Schraw & Moshman, 1995). Along a similar vein, Dunlosky and Tauber (2012), based on the approaches of Nelson and Narens (1990) and Schraw (2002), argued in their theory of *isomechanism* the existence of a general capacity for metamemory, from where they establish that all metacognitive judgments made in different activities, are based on the same cognitive processes. In general, this perspective has been more aligned to researchers who execute their work in the fields of experimental psychology and educational psychology. More recently, this position is recognized in some applications within education, especially in relation to the uses of procedural knowledge. In this regard, for example, Winne and Azevedo (2014) define procedural knowledge as a set of skills that the student has about how to execute cognitive work to perform tasks effectively.

Consistent with this logic, a student can know how to build a first mnemonic letter or how to use a method to optimize an internet search by forming literal strings such as phrases between quotation marks; all of these are procedures and strategies for learning that can be domain specific or general, if they span multiple domains (Winne & Azevedo, 2014; Sawyer, 2014).

Some researchers in the field of metacognition examine differences in domain general vs. domain specific skills. These investigations have included samples from different domains of study such as psychology, education, medicine, and mathematics (Ariel, Dunlosky, & Bailey, 2009; Dunlosky & Rawson, 2012; de Bruin, Dunlosky & Cavalcanti, 2017; Hacker, Bol, Horgan, & Rakow, 2000; Hacker, Bol, & Bahbahani, 2008; Nietfeld, Cao, & Osborne, 2006; Rutherford, 2014). However, many of these studies explore the bases of metacognitive judgments, the capacity of metamemory or the study of the effects of processes of different metacognitive interventions on monitoring accuracy. On the other hand, they do not focus on the analysis of the differences in the metacognitive abilities of knowledge and regulation, based on the hypothetical differences of these components across disciplines or domains. They also do not elucidate the implications for learning these purported differences may bring.

Finally, in a third approach, there are studies from where metacognition is treated as a complex, sophisticated, and hierarchical macroprocess, involving both universal mechanisms more related to general cognition and structural and functional differences within the specific domains of knowledge. These findings have allowed the formulation of new theoretical models of metacognitive monitoring, among which is the proposed *general monitoring* model for metacognitive monitoring (Gutierrez et al., 2016; Schraw et al., 2013, Schraw et al., 2014).

Theoretical Considerations

In this section, theoretical underpinnings to the study of metacognition are discussed. In general, the metacognitive monitoring literature groups different investigations that compare the metacognitive awareness that students have about their actual performance in relation to the expected performance in a criterion task (i.e., confidence in performance judgments). In this way, monitoring accuracy is understood as the degree that individuals judge their performance and how this judgment or belief compares with their actual performance (Nelson, 1996; Gutierrez & Schraw, 2015; Gutierrez & Price, 2017; Gutierrez de Blume, 2017).

There are different measures to determine monitoring accuracy that have been used in previous investigations of metacognitive monitoring. Among them are the estimation of *absolute* accuracy (e.g., G Index or raw frequencies in the 2 x 2 performance/judgment array in Table 1), that evaluates the differences between metacognitive judgments and actual performance, and *relative* accuracy (e.g., Gamma or d') that measures the degree that judgments differentiate or discriminate performance (Serra & Metcalfe, 2009). In turn, the level of agreement between judgments and actual performance has also been evaluated using more sophisticated measures like sensitivity and specificity (Schraw et al., 2013; Schraw et al., 2014; van Stralen et al., 2009).

The prototypical format in metacognitive monitoring studies involves answering a test item and judging whether one's answer is correct or incorrect. Monitoring accuracy and error (bias) are calculated based on different computational formulas that use frequencies of two or more of the four mutually exclusive cells in a 2 x 2 data matrix, a matrix like the one presented below in Table 1.

According to research, cell *a* corresponds to correct performance judged to be correct, cell *b* corresponds to the incorrect performance that is judged to be correct (more commonly referred to as *overconfidence*), cell *c* corresponds to correct performance that it is judged to be incorrect (more commonly known as *underconfidence*), and cell *d* corresponds to incorrect performance that is judged

Table 1

A 2 X 2 Performance-Judgment Data Array for Monitoring Accuracy

Monitoring Judgment	Performance	
	Correct	Incorrect
Correct	a	b
Incorrect	c	d

to be incorrect (Gutierrez et al., 2016; Schraw et al., 2013; Schraw et al., 2014). Different statistical measures combine the four types of information from cells *a* to *d* in various ways to obtain an estimate of the accuracy or error of metacognitive monitoring in a set of test items (locally) or tests as a whole (one holistic global judgment). Thus, cells *a* and *d* correspond to accurate metacognitive monitoring whereas cells *b* and *c* are aligned with erroneous monitoring.

Three metacognitive monitoring models are recognized according to the way that monitoring has been investigated. The first is a classic model derived from the double information flow theory of Nelson and Narens (1990), that typically employs the statistical measure Gamma (Goodman & Kruskal, 1954). In this model, monitoring is considered the link to the available information on the *object* level with the information that the individuals have about their own cognitive resources at a *meta* level. In this way, the information obtained through precise monitoring can be used at the *meta* level to control the subsequent performance of the learner. This model has been called by different researchers as the one factor model, given that its entire research tradition is based on the use of the gamma statistic (Gutierrez et al., 2016; Schraw et al., 2013; Schraw et al., 2014).

A second monitoring model is derived from the medical diagnosis process where the concepts of sensitivity (degree that a test detects disease) and specificity (degree that a test detects the absence of disease) are adapted (Mayer, 2010; Schraw et al.,

2013). Thus, these measures have been employed in studies on metacognitive monitoring in educational settings, such that sensitivity assesses the accuracy of judgments on correct performance, while specificity measures the accuracy of judgments on incorrect performance (Schraw et al., 2013); this monitoring model has been named as two-factor because judgments about correct versus incorrect performance constitute two separate and independent aspects of the metacognitive monitoring process (Gutierrez et al., 2016; Schraw et al., 2013; Schraw et al., 2014).

The third model is referred to as the *general monitoring model*. It assumes that monitoring occurs through two different processes, although inversely related, of accuracy and error, and that individuals obtain metacognitive judgments in different ways. More specifically, the processes related to accurate monitoring judgments are different from those related to erroneous judgments and, as an equally important aspect, judgment errors are not unidimensional but rather divided into discordant judgments in relation to actual performance that lead to overconfidence (e.g., arrogance) and those that lead to underconfidence (e.g., insecurity) (Gutierrez et al., 2016; Schraw et al., 2013; Schraw et al., 2014).

Gutierrez et al. (2016) suggest that metacognitive monitoring can be better understood using this more nuanced *general monitoring model*, where cognitive performance would be the foundation (e.g., performance around specific vocabulary, probability, and paper folding tests). At the same time, this primary and domain specific level can be included in a higher level of monitoring for correct performance and incorrect performance that are assumed to be two different routes to mentally process correct and incorrect metacognitive judgments, regardless of the domain of study. These two previous levels are included in a third level that represents general metacognitive monitoring (Gutierrez et al., 2016; Schraw et al., 2013; Schraw et al., 2014). The general monitoring model is depicted in Figure 1.
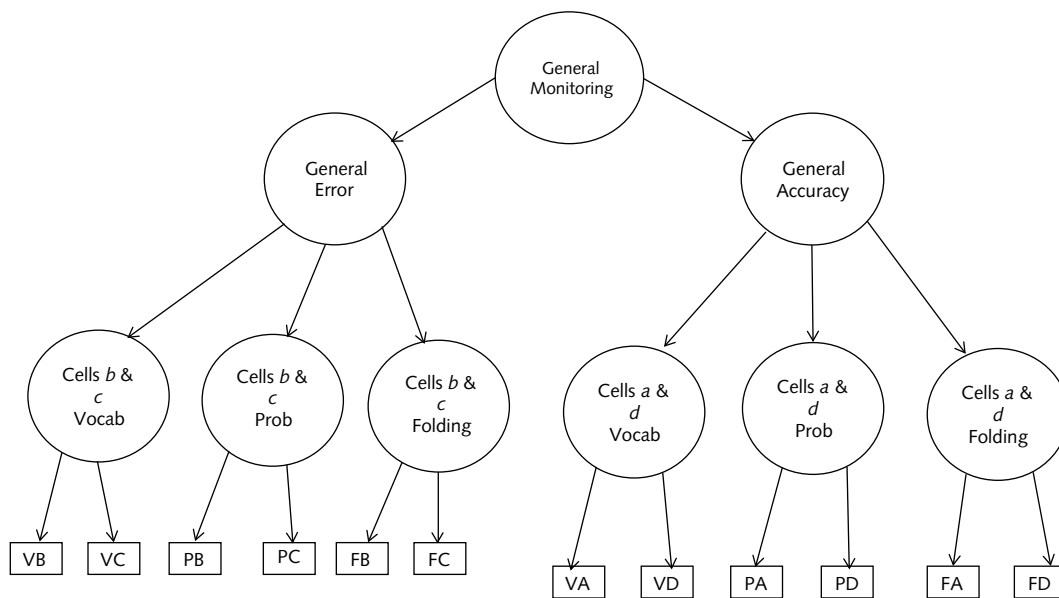
Figure 1. Hypothesized third-order CFA model of the 2 x 2 matrix of raw frequencies for vocabulary, probability, and paper folding tests. The first-order factors represent domain-specific accuracy (i.e., accurate judgments respectively) and domain specific error judgments (i.e., over- and under-confidence). The first letter in each of the manifest variables represents the name of the test (i.e., V = Vocabulary, S = Probability, and F = Paper Folding) and the second letter represents the specific cell in the 2 x 2 matrix.

This notion of a hierarchical framework of metacognitive processes has also been developed from findings proposed in the field of neurophysiology and cognitive neurosciences by different researchers who have studied the anatomo-functional correlates of metacognitive judgments in various domains such as memory and perception processes (e.g., Fleming et al., 2016; Morales et al., 2018). In this regard, Morales et al. (2018), for example, conducted an experiment examining the compromised neural substrates when people made metacognitive judgments during the tasks of perception and memory, combined with performance and stimulus characteristics. When comparing activity patterns of functional magnetic resonance imaging (FMRI) while people evaluated their performance, they found evidence for both domain specific and domain general metacognitive processes. Multi-voxel activity patterns in the anterior prefrontal cortex predicted confidence levels in a domain-specific manner, while domain general signals predicting confidence and accuracy

were found in a generalized network in the front and posterior midline (Morales et al., 2018).

The Present Study

Research question and hypotheses. Based on the reviewed literature , the following research question guided the conduct of this study: What is the effect of domain (in undergraduate programs in education, psychology, and medicine) on the metacognitive skills (knowledge of cognition: declarative, procedural, and conditional; and regulation of cognition: planning, information management, debugging, comprehension monitoring, and evaluation) of undergraduate Colombian students?

Hypotheses: (*1a*) Like previous research (Gutiérrez et al., 2016; Schraw et al., 2014) that supported both the domain general hypothesis and the domain specific hypothesis, in the present study metacognitive awareness was expected to be different among students studying medicine and education and those studying medicine and psychology. However, (*1b*) no significant difference

was expected among undergraduates in education and those in psychology because these domains are more likely to resort to similar metacognitive skill sets compared to the metacognitive skills required in a domain such as medicine, a program of study traditionally recognized as one of the highest regarding academic demand and burden.

Method

Participants and Sample

In the present study, a convenience sampling approach was used. The research involved 507 Colombian university students who during the year 2019 were studying undergraduate programs in education ($n = 156$), psychology ($n = 166$), and medicine ($n = 185$) in two Colombian universities. Of the 507 students, 297 identified as female and 210 identified as male. All students met the following inclusion criteria: age between 20 to 30 years; enrollment as an undergraduate during any of the two semesters of the year 2019; absence of repetition of coursework; and having their signed informed consent for their involvement in the research study. In this way, the distribution of gender and academic program of study is typical of each of the two participating universities.

Materials and Instruments

Metacognitive awareness inventory. The eight components of metacognition (that is, knowledge of cognition: declarative, procedural, and conditional; and regulation of cognition: planning, information management, monitoring, debugging, and evaluation) were measured using the Metacognitive Awareness Inventory (Spanish version; MAI, see Appendix). The MAI was originally developed and validated in English by Schraw and Dennison (1994), and was adapted to be used in Spanish-speakers with Colombian samples by Huertas, Vesga, and Galindo (2014) and by Gutierrez de Blume and Montoya Londoño (2020). The MAI is a 52-item instrument that measures metacognition through its constituent components. Sample items

include: "I constantly wonder if I am meeting my goals" (monitoring); "I try to use strategies that have worked in the past" (procedural knowledge); "I reevaluate what I have learned when I get confused" (debugging strategies); and "I know how well I did in an assessment once the test is over" (evaluation).

The ratings for each item were marked with a vertical bar on a continuous bipolar line of 0-100 (i.e., *not at all true of me* representing 0 and *very true of me* representing 100) on a scale that is 10 cm (i.e. 4 inches) in length. This scoring scheme is superior to an ordinal Likert scale because it improves the reliability of the instrument by increasing the variability of the responses (Schraw & Dennison, 1994; Weaver, 1990). The scores of each participant in the individual scales were obtained by adding all the items that comprise that scale and taking the average. Therefore, each participant had eight composite scores, one for each of the metacognition components.

The Spanish version of the MAI employed in the present study was the version validated by Gutierrez de Blume and Montoya Londoño (2020). Their construct validation study with a sample of 528 undergraduate students demonstrated excellent psychometric properties. For the present study, the eight MAI scales demonstrated appropriate internal consistency reliability coefficients, ranging from .74 to .91 (see Table 2 for descriptive statistics of the sample and internal consistency of each individual scale). This demonstrates that participants provided consistent responses across the various dimensions of metacognition as measured by the MAI (Spanish version), suggesting low measurement error in the hypothesized constructs.

Procedure

Students from the three undergraduate programs included in the study were contacted to solicit their participation. A separate meeting was held for each of the three academic programs during the year 2019, with the collaboration of the instructors in charge of the participants of each

Table 2

Descriptive Statistics for the Sample and Internal Consistency Reliability Coefficients for the Eight mai scales

Scale	M	SD	α
Knowledge of Cognition			
Declarative	67.78	12.82	.74
Procedural	70.59	15.28	.77
Conditional	74.05	13.68	.79
Regulation of Cognition			
Planning			
Information	59.33	14.82	.81
Management	68.32	12.42	.88
Debugging	76.71	14.43	.80
Comprehension	66.28	13.56	.86
Monitoring	62.71	15.38	.91
Evaluation			

Key. M = mean; DS = standard deviation; α = internal consistency reliability coefficient. N=507

of these programs. This provided the space to contact the prospective participants and execute the application of the instrument. In the meeting space with the students, they first learned the specifics of the research, and they read and signed the corresponding informed consent under the advice of one of the instructors. Subsequently, those who expressed interest and agreed to participate in the study, completed the MAI (Spanish version) instrument as a group. The application of the instrument lasted approximately half an hour and the ethical guidelines for studies considered to be of minimum risk with human beings in the country where the data were collected were adhered in the research process (Ministerio de Salud República de Colombia, 1993).

Data Analysis

Before performing data analyses, data were tested for the necessary statistical assumptions and tested for univariate and multivariate outliers. Univariate outliers were reviewed for each group relevant to the present study separately. Inspection of box and whisker plots indicated that, of the original 509 complete cases, two cases were detected that were considered extreme univariate outliers in the education group (standard residuals > 3 standard deviations). Therefore, to avoid biases due to extreme scores that would otherwise

undermine the reliability of the data, these two cases were removed from any subsequent analysis. Next, Mahalanobis Distance and Cook's *D* were used as metrics to evaluate multivariate outliers. None of the 507 remaining cases in the sample were considered multivariate outliers. Therefore, all main analyses continued with these remaining 507 cases with complete data. The data met the assumption of univariate normality (all skewness and kurtosis values were less than the absolute value of 2) and multivariate normality. The data also met the assumption of homogeneity of the error variance (the Levene test *p*-values were all greater than .12) at the univariate level and the homogeneity of the variance-covariance matrices (*p*-value of Box's M=.19) at the multivariate level.

A one-way (domain: education, psychology, medicine) multivariate analysis of variance (MANOVA) was performed. The various metacognitive skills (knowledge of cognition: declarative, procedural, and conditional; and regulation of cognition: planning, information management, monitoring, debugging, and evaluation) served as dependent variables. Bonferroni's adjustment to statistical significance was used to control familywise Type I error rate inflation for univariate omnibus results (.05 / 8=.006) and all post hoc comparisons. All effect sizes for the MANOVA results were reported as η^2. Cohen (1988) specified the following interpretive guidelines for η^2: .010-.059 as small; .060-.139 as medium; and ≥ .140 as large.

Descriptive

Descriptive statistics by group for the eight MAI scales are presented in Table 3, while Tables 4 and 5 present the zero-order bivariate correlations for the sample and by group, respectively.

Evidently, medical students consistently reported significantly lower scores on all eight components of metacognition compared to education and psychology students.

All correlation coefficients were statistically significant at the *p*<.01 level of significance, and all were in the theoretically expected (i.e.,

Table 3

Descriptive Statistics by Group for the Eight MAI scales

Scale	Education (*n* = 156)		Psychology (*n* = 166)		Medicine (*n* = 185)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Knowledge of Cognition						
Declarative	71.16	12.90	70.38	12.92	62.92	11.11
Procedural	71.06	13.81	72.24	14.52	68.80	16.82
Conditional	75.72	11.80	75.13	13.48	71.80	14.92
Regulation of Cognition						
Planning	65.50	13.62	61.90	14.22	52.09	13.22
Information Management	71.24	10.92	67.52	12.87	66.68	12.81
Debugging	79.14	12.52	75.26	14.42	76.09	15.61
Comprehension	70.66	12.77	68.93	13.03	60.47	12.60
Monitoring	66.64	15.17	68.09	13.87	54.87	13.51
Evaluation						

Key. M = mean; DS = standard deviation. N = 507

Table 4

Zero-Order Correlation Matrix of Metacognitive Skills for the Sample and Education Students

Variable	1	2	3	4	5	6	7	8
1. Declarative	-	.65*	.66*	.69*	.73*	.65*	.56*	.65*
2. Procedural	.64*	-	.75*	.55*	.65*	.67*	.58*	.48*
3. Conditional	.72*	.69*	-	.58*	.66*	.73*	.67*	.52*
4. Planning	.69*	.64*	.72*	-	.77*	.61*	.48*	.70*
5. Monitoring	.73*	.69*	.71*	.78*	-	.68*	.59*	.72*
6. Information Management	.64*	.64*	.76*	.64*	.68*	-	.66*	.58*
7. Debugging	.57*	.50*	.70*	.65*	.66*	.66*	-	.41*
8. Evaluation	.69*	.58*	.70*	.74*	.79*	.69*	.65*	-

* p < .01

Note. Correlations above the diagonal are for the sample and those below the diagonal are for the education students.

Table 5

Zero-Order Correlation Matrix of Metacognitive Skills for Psychology and Medical Students

Variable	1	2	3	4	5	6	7	8
1. Declarative	-	.57*	.61*	.56*	.63*	.66*	.58*	.43*
2. Procedural	.74*	-	.76*	.44*	.59*	.71*	.60*	.30*
3. Conditional	.69*	.75*	-	.39*	.59*	.72*	.69*	.29*
4. Planning	.69*	.68*	.70*	-	.65*	.52*	.33*	.54*
5. Monitoring	.75*	.76*	.73*	.76*	-	.65*	.53*	.52*
6. Information Management	.65*	.67*	.72*	.68*	.74*	-	.67*	.42*
7. Debugging	.60*	.63*	.63*	.59*	.70*	.63*	-	.22*
8. Evaluation	.71*	.66*	.66*	.68*	.76*	.74*	.57*	-

* p < .01

Note. Correlations above the diagonal are for medical students and those below the diagonal are for psychology students.

positive) direction. The correlation pattern between the three groups requires a more detailed examination, as they were different depending on domain of study. Interestingly, medical students consistently exhibited weaker correlation coefficients than education and psychology students, suggesting that further exploration of predictive patterns of metacognitive skills and cognitive and performance measures across domains is warranted.

Main Analyses

The MANOVA results revealed that there were statistically significant differences between domains in the linear combination of dependent variables, multivariate $F(16,996)=13.21, p<.001, \eta^2 = .175$. The interpretation of the univariate results is presented below.

After controlling for inflation of the familywise Type I error rate using the Bonferroni adjustment to the p-value for multiple comparisons, the univariate findings were significant between-groups for declarative knowledge, $F(2,504) = 28.31, p < .001, \eta^2 = .101$; conditional knowledge, $F(2,504) = 5.00, p = .006, \eta^2 = .025$; planning, $F(2,504) = 50.05, p < .001, \eta^2 = .166$; monitoring, $F(2,504) = 34.88, p < .001, \eta^2 = .122$; information management, $F(2,504) = 7.24, p = .001, \eta^2 = .028$; and evaluation, $F(2,504) = 49.08, p < .001, \eta^2 = .163$. The univariate results for the procedural knowledge and debugging were not statistically significant, all p values $\geq .03$.

Post hoc pairwise comparisons for statistically significant dependent variables are explained next. The pattern of pairwise differences was relatively consistent in that the mean differences were significant between students of medicine and education and students of medicine and psychology, but not between students of education and psychology. This was the case for declarative knowledge, conditional knowledge, comprehension monitoring, and evaluation. However, for planning and information management, there were also significant differences between education and psychology students, where the former group reported significantly higher mean scores than the latter group. Finally, overall, medical students reported the lowest mean scores on the eight metacognitive components compared to the other two groups (see Table 3 for descriptive statistics by group).

Discussion

All the correlations were statistically significant and all were in the theoretically expected direction, which supports the factorial structure of the classic components of metacognition, focused on knowledge and regulation, drawn from Schraw's original studies to some of his latest research (Gutierrez et al., 2016; Schraw & Dennison, 1994; Schraw, 2002; Schraw et al., 2013; Schraw et al., 2014). In general, statistically significant differences between-groups were found in the case of metacognitive knowledge (declarative and conditional) and regulatory skills (planning, monitoring, information management, and evaluation). However, it should be noted that no significant differences were found in the case of procedural knowledge and debugging skills. These differences were consistent among medical students relative to education and psychology students. Likewise, no differences were found in the metacognitive skills of declarative and conditional knowledge and monitoring and evaluation among education and psychology students. This finding may be due to the way that the Western educational system is focused on the development of language skills, the basis of declarative knowledge, an aspect that seems common in the formation of careers focused on the social sciences, different from what can be presented in health sciences, as is the case of the medical undergraduate program focused on what has been considered as the medical model.

In general, studies on metacognition have not delved into the exploration of differences in metacognitive knowledge and metacognitive regulation processes across domains of study, with a few exceptions. Exceptions include studies regarding specific experimental learning situations in certain disciplinary fields, as those that have addressed, for example, metacognition and its relations with problem solving in mathematics, medicine, and history, at an individual level and in a shared perspective with intelligent tutors (Hurme, Järvelä, Merenluoto & Salonen, 2015; Lajoie, Poitras, Doleck & Jarrell, 2015; Poitras, 2015; Sáiz-Manzanares & Montero-García; 2015), and in the study of the role of metacognitive awareness in listening strategies typical of listening competence in language students (Rahimi & Abedi, 2015). Thus, the investigation of differences of the

central components of metacognition, as in the case of metacognitive knowledge and metacognitive regulation, constitutes a novel endeavor, as the existence of differences in metacognitive skills as a function of domain is shown to be a ripe avenue of investigation.

The present study contributes empirical evidence in the development of this conclusion by providing findings that support the metacognitive differences in specific domains (i.e., education, psychology, and medicine), a finding that is consistent with the metacognitive differences found in studies that used cognitive tasks of vocabulary, probability, and paper folding, previously reported since the initial formulation of the *general monitoring model* (Gutierrez et al., 2016; Schraw & Dennison, 1994; Schraw, 2002; Schraw et al., 2013; Schraw et al., 2014). In this series of studies, Schraw and colleagues examined the extent that metacognitive monitoring is consistent across domains (more commonly known as the *domain-general hypothesis* because metacognitive monitoring is not domain dependent and instead operates as a general set of monitoring skills that transfer and function across different domains) or is domain specific (more commonly known as the *domain-specific* or *encapsulated hypothesis* because metacognitive monitoring skills are domain specific and, therefore, do not transfer or function similarly across domains).

In their last two investigations, Schraw and his associates posit the argument that perhaps metacognitive monitoring is a combination of domain-specific and domain-general hypotheses (Gutiérrez et al., 2016; Schraw et al., 2014). They argue that metacognitive monitoring begins as a domain-specific skill early in development and progressively becomes more general in adolescence and, especially, in adulthood. This conjecture is certainly consistent with modern conceptualizations of the theory of self-regulated learning (Bandura, 2006; Boekaerts, 1999; Zimmerman, 2000). In the present study, medical students reported lower scores in the eight components of metacognition

evaluated compared with students from other undergraduate programs such as education and psychology. This finding is interesting, given the high academic demand that training in the field of study of medicine represents in any country in the world. Medicine is a discipline of knowledge that some researchers recognize as a domain where students can demonstrate exhaustion as a result of the arduous demands of this domain (e.g., as a result of the stress related to academic work, the handling of a large volume of knowledge and information, the excess of tasks, as well as the exposure to the suffering of patients, which represents a very high cognitive and emotional demand), that may trigger feelings of incompetence in the student, with important consequences for learning (Dyrbye, et al., 2014; Phinder-Puente et al., 2014).

Like Schraw and colleagues (Gutiérrez et al., 2016; Schraw et al., 2014), this study also found support for both the domain-general hypothesis and the domain-specific hypothesis, albeit in more nuanced ways. Presumably, as social sciences, psychology and education share much more in common with each other than with medicine, and therefore metacognitive skills can be more easily transferred across those domains than in medicine. Although it has been documented worldwide that, in general, students with high intellectual capacity are admitted to study in medical programs, only a very small percentage of students are actually admitted. Once medical students begin their studies, their learning and performance possibilities vary widely, between profiles of students who achieve in-depth learning and high levels of competence in relation to their training, and those students who barely manage to pass and even experience academic decline and, ultimately, dropout (Abdulghani, 2009; Abdulghani, et al., 2014; Arulampalam, Naylor, & Smith, 2004).

The lower mean scores on all of the components of metacognition in medical students compared to the other two groups seem to demonstrate the need to incorporate the intentional and explicit teaching of metacognitive skills in undergraduate

careers with such high academic demands, as would be the case for medical programs, to promote in medical students a more reflective attitude towards their own learning. This approach favors the development of a greater metacognitive awareness of the possibilities of knowing oneself well during learning episodes and to control the process of self-regulation of learning. In this regard, some researchers have argued that, although much of medical education involves immersion in clinical learning environments, students rarely, if at all, receive explicit instructions in their training on how to manage their own learning. (Aukes, et al., 2007; Westberg, & Jason, 1994). In fact, some researchers have described the culture of teaching medicine as influenced by a formative action that can sometimes be unreflective, given that long hours of work and the burden of clinical service can limit the time practicing doctors have to reflect on their experience, an aspect that, at the same time, contributes to limiting the possibilities of time invested in training in the perspective of a new form of patient care, based on a process of constant self-evaluation and lifelong learning by the doctor in training (Aukes et al., 2007; Nothnagle, Goldman, Quirk, & Reis, 2010; Westberg, & Jason, 1994).

In this sense, Quirk (2006) posits that in the face of the challenges of medical training, medical educators should explicitly teach medical students metacognitive skills that allow them to: (a) define and prioritize their goals compared to their study; (b) anticipate and assess their specific learning needs relative to goals; (c) organize (and reorganize) their experiences to meet their learning needs; (d) define their own perspective and recognize differences in the perspectives of others; and (e) continuously monitor their learning, to exercise control over their knowledge base, the resolution of problems and interactions with others, as well as against their own process of metacognitive reflection. Presumably, this should benefit their learning because it is predicated in real clinical practice with the ability to act quickly and decisively (Quirk, 2006; 2014).

It is logical that higher scores are evident in the eight components of metacognition for students in undergraduate programs in psychology and education, as traditionally these are careers where training spaces focused on reflection on self, on practice and in general, and on introspection are provided in abundance, as they are disciplines focused on reflective thinking (Cooper, & Wieckowski, 2017; Foong, Nor, & Nolan, 2018). In general, the development of reflective thinking, as the basis of metacognitive action, requires training spaces to make sense of the experience in relation to oneself, with others, with contextual conditions, and with the development of the ability to monitor and plan future experiences, as well as to develop a greater awareness about learning (Fullana et al., 2016; Ryan, 2013).

Finally, differences were also found for planning and information management skills insofar as education students reported significantly higher scores than psychology students in these measures, a finding that can plausibly be explained by the emphasis these skills are given in education programs. Preservice teachers, for instance, are trained from the very beginning on skills such as defining teaching objectives, the search for sources, planning, the establishment of didactic sequences, monitoring of the state of their students' learning, and forecasting the evaluation process (Jiang et al., 2016; Stewart et al., 2007). This is done while education students must plan and organize the sources and resources for their teaching, skills that are required based on their practice processes, assumed as reflective practice, that according to their particular undergraduate education specialization, they can start in tandem with their academic preparation in the basic and initial training cycle. In this regard, some researchers have described that reflective practice has been included in initial teacher training programs as part of the international movement to reform teaching and improve the quality of education

since the 1980s (Adams & Mabusela, 2014; Jiang et al., 2016; Lee, Irving, Pape, & Owens, 2015; OCDE, 1989; Robinson, Anderson-Harper, & Kochan, 2001). Currently, it is considered as a compulsory competency in many preservice teacher training programs (Collin, Karsenti, & Komis, 2013; Richardson, 1990).

Implications for Theory, Research and Practice

The findings of the present study, although based on self-report measures, tentatively have a significant relation with the theory of self-regulated learning and the *general monitoring* framework proposed by Gutiérrez et al. (2016). More specifically, they indicate that, insofar as the theoretical statements are consistent depending on the task and domain, this relation is much more complex than in initial hypotheses. The extent that these theories are generalized across all domains, are domain-specific, or a combination of both hypotheses, should be examined more fully by researchers. Therefore, research efforts should consider recruiting samples from a variety of different domains (e.g., physics, biology, engineering, fine arts, etc.) and different roles within those domains (e.g., clinician, teacher, researcher) to better understand to what extent metacognitive skills are domain-general or domain-specific. Finally, regarding practice, the findings suggest that additional applied experiments should be designed, examining the extent that interventions aimed at improving metacognitive skills can be generalized across different learning domains.

Avenues for Future Research

The findings of the present study found partial support for the *domain-general hypothesis* and the *domain-specific hypothesis*. Evidently, the results for education and psychology students (domain-general) were much more consistent than those found for medical students (domain-specific). However, the present study used only a self-report measure, albeit supported by appropriate

psychometric properties from a wide range of languages and cultures. Future research should also employ cognitive or performance measures like those implemented by Gutierrez et al. (2016) and Schraw et al. (2014) and explore how these work in various domains under study. An interesting approach that would further support both hypotheses would be studies investigating block domains (i.e., social sciences like education and psychology compared to engineering or biology) to more fully determine the specificity and/or generality of metacognitive skills.

Likewise, it would be beneficial to investigate whether medical students who aspire to continue training in areas of clinical specialties (for example, family general practitioners, pediatricians, etc.) differ in their metacognitive ability compared to students who choose to become researchers (for example, epidemiologists, immunologists, etc.). Finally, additional research on the cross-cultural implications of our findings would also be useful.

Methodological Reflections and Limitations

It is important to be transparent and highlight the limitations of this research. First, the study employed only a self-report measure that, although considered the gold standard for measuring the construct and has been validated in different languages and cultures, including samples from the Colombian population, is only one measure. Second, it could be that, due to phenomena such as social desirability bias, participants were not completely open and honest about their true metacognitive abilities, a limitation that is most evident on self-report measures. Finally, some of the effect sizes reported in our MANOVA are considered small effects. Therefore, our findings and conclusions may not fully or adequately capture the true effects in the populations under study.

However, despite these limitations, it is important to highlight some strengths of the study. It employed a robust sample size, and thus, increased the stability of the findings (i.e., they are not likely to be spurious). Likewise, the study was implemented

in the context of real classrooms, where these students carry out their learning process, instead of the artificial environment of a laboratory and, therefore, it is more ecologically valid. In sum, despite the limitations, the present study contributes to a better understanding of how metacognitive skills operate according to different learning domains.

Conclusion

The results found in the present investigation contribute to the empirical and, in some way, cross-cultural validation of the multilevel metacognitive *general monitoring model*, to the extent that they support the perspective that metacognitive monitoring is considered to be a combination of the domain-specific hypothesis and the domain-general hypothesis (Gutierrez et al., 2016; Schraw et al., 2014). In particular, the present study is of special relevance to the authors, as it constitutes the first research published in Spanish on the empirical support for the *general monitoring model*. It is certainly worth noting that psychology and education students (in support of the domain-general view of metacognitive monitoring) reported similar mean scores on most of the eight metacognitive skills components. Equally as important, medical students reported the lowest mean scores overall across all eight components of metacognition. The fact that our sample size was robust lends credence to the findings of the present study, as they are not likely due to chance. It is hoped that the present study spurs additional debate and future research on the *general monitoring model* and the examination of the *domain-general* versus *domain-specific* hypotheses.

References

Abdulghani, H. M. (2009). Admission criteria for Saudi Health Colleges: The current status and a literature review. *Medical Channel, 15*, 18-21.

Abdulghani, H., Al-Drees, A., Khalil, M., Ahmad, F., Ponnamperuma, G., & Amin, Z. (2014). What factors determine academic achievement in high achieving undergraduate medical students? A qualitative study. *Medical Teacher, 36*, 43-48. https://doi.org/10.3109/0142159x.2014.886011

Adams, J., & Mabusela, M. (2014). A metacognitive approach to teacher development: Supporting national professional diploma in education students. *Mediterranean Journal of Social Sciences, 5*, 289-296. https://doi.org/10.5901/mjss.2014.v5n15p289

Ariel, R., Dunlosky, J., & Bailey, B. (2009). Agenda-based regulation of study-time allocation: When agendas override item-based monitoring. *Journal of Experimental Psychology: General, 138*, 432-447. https://doi.org/10.1037/a0015928

Arulampalam, W., Naylor, R., & Smith, J. (2004). Factors affecting the probability of first year medical student dropout in the UK: A logistic analysis for the intake cohorts of 1980–92. *Medical Education, 38*, 492-503. https://doi.org/10.1046/j.1365-2929.2004.01815.x

Aukes L. C., Geertsma, J., Cohen-Schotanus, J., Zwierstra, R.P., & Slaets, J. P. (2007). The development of a scale to measure personal reflection in medical practice and education. *Medical Teacher, 29*, 177-182. https://doi.org/10.1046/ht10.1080/01421590701299272

Bandura, A. (2006). Toward a psychology of human agency. *Perspectives on Psychological Science, 1*, 164-180. https://doi.org/10.1046/10.1111%2Fj.1745-6916.2006.00011.x

Boekaerts, M. (1999). Self-regulated learning: Where we are today. *International Journal of Educational Research, 31*, 445-457. https://doi.org/10.1046/10.1016/S0883-0355(99)00014-2

Collin, S., Karsenti, T., & Komis, V. (2013). Reflective practice in initial teacher training: critiques and perspectives. *Reflective Practice, 14*, 104-117. https://doi.org/10.1046/10.1080/14623943.2012.732935

Cooper, L. D., & Wieckowski, A. T. (2017). A structured approach to reflective practice training in a clinical practicum. *Training and Education in Professional Psychology, 11*, 252-259. https://doi.org/10.1046/10.1037/tep0000170

de Bruin, A., Dunlosky, J., & Cavalcanti, R. (2017). Monitoring and regulation of learning in medical education: The need for predictive cues. *Medical Education, 51*, 575-584. https://doi.org/10.1046/10.1111/medu.13267

Dunlosky, J., & Rawson, K. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction, 22*, 271-280. https://doi.org/10.1016/j.learninstruc.2011.08.003

Dunlosky, J., & Tauber, S. K. (2012). Understanding people's metacognitive judgments: An isomechanism framework and its implications for applied and theoretical research. In T. Perfect & S. Lindsay (Eds.), *The SAGE handbook of applied memory* (pp. 1-10). London: Sage Publications Inc.

Dyrbye, L., West, C., Satele, D., Boone, S., Tan, L., Sloan, J., & Shanafelt, T. (2014). Burnout among U.S. medical students, residents, and early career physicians relative to the general U.S. population. *Academic Medicine, 89*, 443-451. https://doi.org/10.1097/acm.0000000000000134

Fleming, S., Massoni, S., Gajdos, T., & Vergnaud, J. (2016). Metacognition about the past and future: Quantifying common and distinct influences on prospective and retrospective judgments of self-performance. *Neuroscience of Consciousness, 2016*, 1-12. https://doi.org/10.1093/nc/niw018

Foong, L., Nor, M., & Nolan, N. (2018). The influence of practicum supervisors' facilitation styles on student teachers' reflective thinking during collective reflection. *Reflective Practice, 19*, 225-242. https://doi.org/10.1080/14623943.2018.1437406

Fullana, J., Pallisera, M., Colomer, J., Fernández, R., & Pérez-Burriel, M. (2016). Reflective learning in higher education: a qualitative study on students' perceptions. *Studies in Higher Education, 41*, 1008-1022. https://doi.org/10.1080/03075079.2014.950563

Garner, R. (1987). *Metacognition and reading comprehension*. United States of America.: Ablex Publishing.

Goodman, L. A., & Kruskal, W. H. (1954). Measures of association for cross classification. *Journal of the American Statistical Association, 49*, 732-764. https://doi.org/10.1080/01621459.1954.10501231

Gutierrez, A. P., & Schraw, G. (2015). Effects of strategy training and incentives on students' performance, confidence, and calibration. *Journal of Experimental Education, 83*, 386–404. https://doi.org/10.1080/00220973.2014.907230

Gutierrez, A. P., Schraw, G., Kuch, F., & Richmond, A. S. (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. P., & Price, A. F. (2017). Calibration between undergraduate students' prediction of and actual performance: The role of gender and performance attributions. *The Journal of Experimental Education, 85*, 486-500. https://doi.org/10.1080/00220973.2016.1180278

Gutierrrez De Blume, A. P. (2017). The effects of strategy training and an extrinsic incentive on fourth-and fifth-grade students' performance, confidence, and calibration accuracy. *Cogen Education, 4*(1), 1-17. https://doi.org/10.1080/2331186X.2017.1314652

Gutierrez de Blume, A. P., & Montoya Londoño, D. M. (2020). *Validation of the MAI in Colombian university students*. [Manuscript submitted for publication.]

Hacker, D. J., Bol, L., Horgan, D. D., & Rakow, E. A. (2000). Test prediction and performance in a classroom context. *Journal of Educational Psychology, 92*, 160-170. https://doi.org/10.1037/0022-0663.92.1.160

Hacker, D, Bol, L., & Bahbahani, K. (2008). Explaining calibration accuracy in classroom contexts: The effects of incentives, reflection, and explanatory style. *Metacognition and Learning, 3*, 101-121. https://doi.org/10.1007%2Fs11409-008-9021-5

Huertas, A., Vesga, G., & Galindo, M. (2014). Validación del instrumento inventario de habilidades metacognitivas "MAI" con estudiantes colombianos. *Revista Praxis & Saber, 5*, 55-74. https://doi.org/10.19053/22160159.3022

Hurme, T. R., Järvelä, S., Merenluoto, K., & Salonen, P. (2015). What makes metacognition as socially shared in mathematical problem solving. In A. Peña-Ayala (Ed.), *Metacognition: Fundaments, applications and trends. A prolife of the current state -of-the-art* (pp. 259-275). New York: Springer.

Jiang, Y., Ma, L., & Gao, L. (2016). Assessing teachers' metacognition in teaching: The Teacher Metacognition Inventory. *Teaching and Teacher Education, 59*, 403-413. https://doi.org/10.1016/j.tate.2016.07.014

Lajoie, S., Poitras, E., Doleck, T., & Jarrell, A. (2015). Modeling metacognitive activities in medical

problem-solving with BioWorld. In A. Peña-Ayala (Ed.), *Metacognition: Fundaments, applications and trends. A prolife of the current state -of-the-art* (pp. 323-341). New York: Springer.

Lee, S., Irving, K., Pape, S., & Owens, D. (2015). Teachers' use of interactive technology to enhance students' metacognition: Awareness of student learning and feedback. *Journal of Computers in Mathematics & Science Teaching, 34*, 175-198. Retrieved from https://www.learntechlib.org/primary/p/114753/

Mayer, D. (2010). *Essential evidence-based medicine* (2nd ed.). Cambridge, England: Cambridge University Press.

Ministerio de Salud República de Colombia. (1993). Resolución 008430 del 4 de octubre. Recuperada el 15 de enero de 2020 de https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/DE/DIJ/RESOLUCION-8430-DE-1993.PDF

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

Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist, 51*, 102-116. https://doi.org/10.1037/0003-066X.51.2.102

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

Nietfeld, J. L., Cao, L., & Osborne, J. W. (2006). The effect of distributed monitoring exercises and feedback on performance, monitoring accuracy, and self-efficacy. *Metacognition and Learning, 1*, 159–179. https://doi.org/10.1007/s10409-006-9595-6

Nothnagle, M. D., Goldman, R., Quirk, M., & Reis, S. (2010). Promoting self-directed learning skills in residency: A case study in program development. *Academic Medicine, 85*, 1874-1879. https://doi.org/10.1097/acm.0b013e3181fa02a4

Organisation de Coopération et de Développement Economiques (OCDE). (1989). *Les écoles et la qualité: un rapport international*. Paris: OCDE.

Phinder-Puente, M., Sánchez-Cardel, A., Romero- Caste-llanos, F., Vizcarra-García, J., & Sánchez-Valdivieso, E. (2014). Percepción sobre factores estresantes en estudiantes de Medicina de primer semestre, sus padres y sus maestros. *Investigación en Educación Médica, 3*,139-146. https://doi.org/10.1016/S2007-5057(14)72740-1

Poitras, E. (2015). Modeling metacognitive activities in medical problem-solving with BioWorld. In A. Peña-Ayala (Ed.), *Metacognition: Fundaments, applications and trends. A prolife of the current state -of-the-art* (pp. 345-367). New York: Springer.

Quirk, M. (2006). *Intuition and metacognition in medical education. Keys to developing* expertise. New York: Springer.

Quirk M. (2014) The Metacognitive Competency: The Key to Lifelong Learning. In: Kalet A., Chou C. (eds) *Remediation in Medical Education*. New York: Springer.

Rahimi, M., & Abedi, S. (2015). The role of metacognitive awareness of listening strategies in listening proficiency: The case of language learners with different levels of academic self-regulation. In A. Peña-Ayala (Ed.), *Metacognition: Fundaments, applications, and trends. A prolife of the current state -of-the-art* (pp. 169-189). New York: Springer.

Richardson, V. (1990). The evolution of reflective teaching and teacher education. In R. Clift, R. Houston, & M. Pugash (Eds.), *Encouraging reflective practice in education: An analysis of issues and programs* (pp. 3-20). New York: Teachers College Press.

Robinson, E., Anderson-Harper, H., & Kochan, F. (2001). Strategies to improve reflective teaching. *Journal of Pharmacy Teaching, 8*, 49-58. https://doi.org/10.1300/J060v08n04_04

Rutherford, T. (2014). Calibration of confidence judgments in elementary mathematics: measurement, development, and improvement. (Tesis de doctorado). University of California, Irvine, Estados Unidos. Recuperada de https://escholarship.org/content/qt99z17038/qt99z17038.pdf

Ryan, M. (2013). The pedagogical balancing act: Teaching reflection in higher education. *Teaching in Higher*

Education, 18, 144-155. https://doi.org/10.1080/1356 2517.2012.694104

Sáiz-Manzanares, M., & Montero-García, E. (2015). Metacognition, self-regulation and assessment in problem-solving processes at university. In A. Peña-Ayala (Ed.), *Metacognition: Fundaments, applications and trends. A prolife of the current state -of-the-art* (pp. 107-130). New York: Springer.

Sawyer, K. (2014). *The Cambridge handbook of the learning sciences*. New York: Cambridge University Press.

Schraw, G. (2002). Promoting general metacognitive awareness. In H. Hartman (Ed.), *Metacognition in learning and instruction: Theory, research, and practice* (pp. 3-16). London: Kluwer Academic.

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

Schraw, G., & Moshman, D. (1995). Metacognitive theories. *Educational Psychology Review*, *7*(4), 351-371. https://doi.org/10.1007/BF02212307

Schraw, G., Kuch, F., & Gutierrez, G. (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, A. S. (2014). Exploring a three-level model of calibration accuracy. *Journal of Educational Psychology*, *106*, 1192-1202. https://doi.org/10.1037/a0036653

Serra, M. J., & Metcalfe, J. (2009). Effective implementation of metacognition. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 278-298). Mahwah, NJ: Erlbaum.

Stewart, P., Cooper. S., & Moulding, L. (2007). Metacognitive development in professional educators. *The Researcher, 21*, 32-40. Retrieved from http://www.nrmera.org/wp-content/uploads/2016/02/Researcherv21n1Stewart.pdf

van Stralen, K. J., Stel, V. S., Reitsma, J. B., Dekker, F. W., Zoccali, C., & Jager, K. J. (2009). Diagnostic methods I: Sensitivity, specificity, and other measures of accuracy. *Kidney International, 75*, 1257-1263.

Weaver, C. (1990). Constraining factors in calibration of comprehension. *Journal of Experimental Psychology:*

Learning, Memory, and Cognition, *16*, 214-222. https://doi.org/10.1037/0278-7393.16.2.214

Westberg J., & Jason H. (1994). Fostering learners' reflection and self-assessment. *Family Medicine, 26*, 278-282.

Winne, P., & Azevedo, R. (2014). Metacognition. In K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 63-87). New York: Cambridge University Press

Winne, P., & Marzouk, Z. (2019). Learning strategies and self -regulated learning. In J. Dunlosky y K. A. Rawson (Eds.), *The Cambridge handbook cognition and education* (pp. 606-715). Cambridge: Cambridge University Press.

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13-39). San Diego, CA: Academic Press.

Appendix
MAI

Test de Conciencia Metacognitiva General -(Schraw & Dennison, 1994).

Nombre: _____

Edad: _____**Programa académico:**_____ **semestre:**_____

Indicaciones: Por favor dibuje una línea vertical sobre la línea continua en cada uno de los enunciados, que mejor corresponda a qué tan cierto es cada una de las siguientes preguntas sobre usted.

Para nada cierto sobre mí Muy cierto sobre mí

```
  0  5  10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100
  0%                                                        100%
```
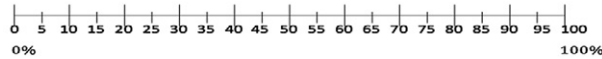
Por ejemplo: Entre más cerca este la línea de "Para nada cierto sobre mí", menos cierto es el enunciado sobre usted, por el contrario, entre más cerca este la línea a "muy cierto sobre mí", más verdadero será el enunciado sobre usted. De la misma manera, dibujar la línea muy al comienzo o muy al final del enunciado (0 – 100), indica que el enunciado es muy cierto (100) o falso (0) con respecto a usted.

1. Constantemente me pregunto si estoy cumpliendo mis metas.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 / 0% ... 100%
2. Considero varias opciones con respecto a un problema antes de contestar.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 / 0% ... 100%
3. Intento utilizar estrategias que han funcionado en el pasado.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 / 0% ... 100%
4. Me organizo mientras aprendo, de tal manera que tenga tiempo suficiente.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 / 0% ... 100%
5. Tengo claras cuáles son mis fortalezas y debilidades intelectuales.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 / 0% ... 100%
6. Siempre pienso en lo que en realidad necesito aprender antes de comenzar una tarea.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 / 0% ... 100%
7. Sé qué tan bien me fue en una evaluación, una vez termine la prueba.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 / 0% ... 100%
8. Establezco metas específicas antes de comenzar una tarea.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 / 0% ... 100%
9. Disminuyo mi ritmo de trabajo cuando encuentro información importante.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 / 0% ... 100%
10. Sé cuál es la información más importante que debo aprender.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 / 0% ... 100%
11. Me pregunto si he tenido en cuenta todas las opciones cuando tengo que resolver un problema.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 / 0% ... 100%
12. Soy bueno organizando información.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 / 0% ... 100%
13. Conscientemente enfoco mi atención en la información importante.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 / 0% ... 100%
14. Tengo un propósito específico con cada una de las estrategias que utilizo.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 / 0% ... 100%
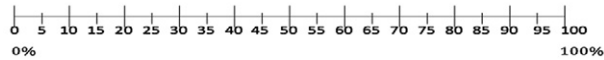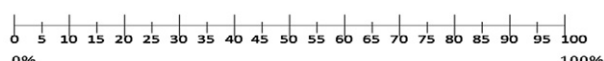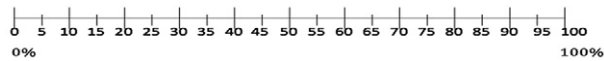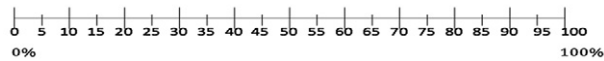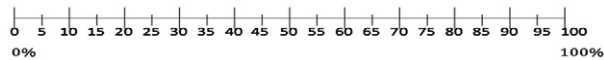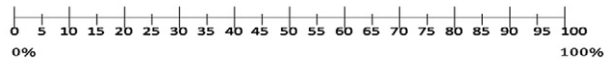
15. Aprendo mejor cuando sé algo con respecto al tema.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 0% 100%
16. Sé lo que el profesor espera que yo aprenda.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 0% 100%
17. Soy bueno recordando información.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 0% 100%
18. Utilizo diferentes estrategias de aprendizaje dependiendo de la situación.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 0% 100%
19. Después de terminar una tarea, me pregunto si había una forma más fácil de resolverla.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 0% 100%
20. Tengo control sobre qué tan bien aprendo.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 0% 100%
21. Periódicamente estoy estudiando para ayudarme a comprender relaciones importantes.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 0% 100%
22. Me hago preguntas acerca de las lecturas antes de comenzar a leer.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 0% 100%
23. Pienso en varias formas de resolver un problema y selecciono la mejor.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 0% 100%
24. Hago resúmenes de lo que he aprendido una vez termino.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 0% 100%
25. Pido la ayuda de otros cuando no comprendo algo.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 0% 100%
26. Puedo motivarme a aprender lo que necesito aprender.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 0% 100%
27. Soy consciente de que estrategias debo utilizar cuando estudio.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 0% 100%
28. Puedo analizar la utilidad de las estrategias que uso cuando estudio.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 0% 100%
29. Utilizo mis fortalezas intelectuales para compensar mis debilidades.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 0% 100%
30. Me enfoco en el significado y significancia de la información novedosa.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 0% 100%
31. Puedo generar mis propios ejemplos para que la información sea más significativa.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 0% 100%
32. Puedo juzgar muy bien que tan bien comprendo una temática o tema.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 0% 100%
33. Por lo general, utilizo estrategias de aprendizaje útiles automáticamente.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 0% 100%
34. Por lo general, puedo disminuir mi ritmo de trabajo para saber si estoy comprendiendo.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 0% 100%
35. Sé cuándo las estrategias que utilizo serán más efectivas.	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 0% 100%

36. Puedo saber que tan bien he logrado mis metas una vez he terminado.

0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100
0%   100%

37. Realizo mapas conceptuales para ayudarme a comprender mientras estudio.

0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100
0%   100%

38. Me pregunto si he tenido en cuenta todas las opciones antes de resolver un problema.

0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100
0%   100%

39. Trato de poner toda la información en mis propias palabras.

0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100
0%   100%

40. Cambio las estrategias cuando no logro comprender muy bien.

0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100
0%   100%

41. Utilizo la estructura organizacional del texto para comprender mejor.

0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100
0%   100%

42. Leo las instrucciones cuidadosamente antes de comenzar una tarea.

0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100
0%   100%

43. Me pregunto si lo que estoy leyendo está relacionado con lo que ya se.

0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100
0%   100%

44. Reevalúo lo que he aprendido cuando me confundo.

0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100
0%   100%

45. Organizo mi tiempo para lograr todas mis metas.

0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100
0%   100%

46. Aprendo más cuando estoy interesado en el tema.

0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100
0%   100%

47. Intento estudiar por partes para tener una mejor comprensión.

0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100
0%   100%

48. Me enfoco en los significados generales, más que en los específicos.

0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100
0%   100%

49. Me hago preguntas con respecto a que tan bien estoy haciendo las cosas, cuando aprendo nueva información.

0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100
0%   100%

50. Me pregunto si aprendí tanto como debería, una vez termino la tarea.

0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100
0%   100%

51. Me detengo y puedo volver a revisar información que aún no me es clara.

0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100
0%   100%

52. Me detengo y vuelvo a leer cuando estoy confundido.

0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100
0%   100%