Metacognition and Learning
https://doi.org/10.1007/s11409-024-09388-9

Psychometric properties of the Metacognitive Awareness Inventory (MAI): standardization to an international spanish with 12 countries

Antonio P. Gutierrez de Blume[1] · Diana Marcela Montoya Londoño[2] · Virginia Jiménez Rodríguez[3] · Olivia Morán Núñez[4] · Ariel Cuadro[5] · Lilián Daset[5] · Mauricio Molina Delgado[6] · Claudia García de la Cadena[7] · María Beatríz Beltrán Navarro[8] · Aníbal Puente Ferreras[3] · Sebastián Urquijo[9] · Walter Lizandro Arias[10]

Received: 9 January 2024 / Accepted: 26 May 2024

Abstract
Metacognition is defined as a higher-order thinking skill that enables individuals to monitor, control, and regulate their thinking and behavior. In education, this skill is important, as learners need to self-regulate their learning behaviors for successful lifelong learning. Thus, it is essential for educators and learners alike to know their metacognitive skills. Researchers can assist in this endeavor by developing sound and valid quantitative measures for psychological phenomena such as metacognition. No measure is more commonly used for this purpose than the Metacognitive Awareness Inventory (MAI). In the present study, the International Group on Metacognition validated the MAI employing a standard, international Spanish with a robust sample of 12 Spanish-speaking countries and 1,622 undergraduate university students. Results revealed a solid final baseline confirmatory factor analysis model for all 12 countries that supports the original two-factor structure reported in English-speaking samples from the United States. Additionally, multigroup measurement invariance analyses revealed that although five parameters varied slightly across some countries, chi-square difference tests indicated that the comparison model with these constraints freely estimated was not significantly better than the fully constrained null model, supporting measurement invariance across countries. Thus, our version of the MAI using standard, international Spanish is a valid and reliable tool for measuring metacognitive awareness in Spanish-speaking countries.

Keywords International spanish validation · Metacognition · Subjective metacognitive awareness · Self-regulated learning

Introduction

Metacognition has been generally defined as a higher-order thinking skill that involves monitoring and control of one's cognitive resources, and it is recognized as a major component of self-regulated learning (Dinsmore et al., 2008; Efklides, 2006, 2008;

Extended author information available on the last page of the article

Gutierrez & Schraw, 2015; Gutierrez de Blume et al., 2020; Schraw & Dennison, 1994; Schraw, 1998; Winne & Azevedo, 2014). Schraw and Dennison (1994) were the first to introduce the notion of metacognitive awareness (i.e., the Metacognitive Awareness Inventory [MAI]), a specific metacognitive phenomenon, which they defined as individuals' perceptions of their own metacognitive skills. In their seminal study, they developed the first self-report instrument of metacognitive awareness with a sample of 197 university undergraduate students. The 52-item measure was found to be comprised of two higher-order factors of knowledge of cognition (incorporating declarative, procedural, and conditional knowledge) and regulation of cognition (subsuming planning, monitoring, information management, debugging, and evaluation). While the MAI has a long history of being translated and employed in various cultures and languages, previous research has only tested the viability of the measure using single samples from their respective populations. Nevertheless, researchers have assumed that the results of these studies generalized to the entire population without empirically testing this claim. To address this research gap, the present study employed the Spanish-version of the MAI that was developed using an international Spanish among a sample of 1,622 university undergraduate students from 12 Spanish-speaking countries in Latin America and Spain.

Background

Different researchers propose that studies on metacognition constitute one of the most promising alternatives to generate a change in training processes, to the extent that their promotion in the classroom favors a higher level of self-awareness of students about their own learning processes and cognitive resources. Metacognition: 1) allows one to know oneself and reflect on one's own capabilities and preferences in the face of learning demands; 2) allows one to identify possible difficulties and problems as well as select the best strategies to solve them; and 3) monitor and employ evaluation and planning processes of learning behavior, in a perspective of self-regulation of learning, which favors the development of one's own agency (Cheng & Chan, 2021; Pandey & Mohan, 2023; Schraw, 2009a; Veenman et al., 2014). In fact, researchers have argued that a person's metacognitive performance is the greatest predictor of their learning potential (Veenman et al., 2006; Wang et al., 1990). The development of the capacity for agency is based on the idea that it allows us to recognize that people are capable of making metacognitive evaluations about the control they have over the results of their own actions and about when to execute them, and is oriented towards the capacity of exercising control over the nature of one's own quality of life; in this case, associated with learning behavior (Bandura, 2001; Metcalfe & Greene, 2007).

From this perspective, the development and evaluation of metacognitive abilities is essential to face the demands of the development of contemporary thinking. This is often characterized by the need to train the person in the skills necessary to acquire new, complex knowledge and to build and access deep learning processes, given the changing dynamics of the global labor market. In this context, professions and jobs increasingly involve evolving roles that require problem-solving skills, thinking independent and creative professionals, as well as development of skills to enhance metacognitive thinking and self-regulated learning (Double & Birney, 2019; Dunlosky & Rawson, 2019; Ozturk, 2017; Pintrich et al., 2000; Quinn et al., 2021; Winne & Perry, 2005; Zohar & Dori, 2012, Zohar, 2020).

In general, metacognition is considered a higher-order executive thinking process (Flores-Lázaro et al., 2014a, 2014b; Flores et al., 2008; Follmer & Sperling, 2016;

Livingston, 2003; Veenman et al., 2006), which involves knowledge and regulation of the activity of cognitive processes and products (Flavell, 1976). Likewise, it represents a critical, reflective thinking process that shows the level of consciousness that people reach in relation to their own learning. This involves understanding and manipulating one's own cognitive abilities to perform more efficiently and effectively (Schraw & Dennison, 1994; Serra & Metcalfe, 2009; Zohar, 2020).

Metacognitive knowledge is understood as everything that the individuals know about themselves as learners (declarative knowledge), everything they know about how they can do it, which includes the use of learning strategies (procedural knowledge), and about when, where, and why to use the knowledge that tasks require (conditional or strategic knowledge), and it is considered the basis of individuals' regulation processes (Brown, 1987; Cross & Paris, 1988; Garner, 1990; Schraw & Dennison, 1994; Ozturk, 2017; Pintrich et al., 2000; Schraw, 1998, 2002; Veenman et al., 2006; Zohar, 2020).

On the other hand, metacognitive regulation refers to the way in which individuals control their cognitive activity; therefore, it involves the type of practical strategies that individuals use to regulate and monitor their own resources and processes, as well as the results of their learning. Regulatory skills are: 1) planning (which involves anticipation, setting objectives, choosing strategies, designing a sequence of steps, allocation of time and resources, etc.); 2) follow-up, also called monitoring skill (involves online supervision of the task in progress, which requires processes such as testing, rehearsal, review, adjustment, and awareness of the state of performance); and, finally, 3) evaluation (which can be understood as a value judgment regarding one's own performance, which involves the comparison between the objectives and the result, and the analysis of the use of the strategies and their level of effectiveness, both the process and the product obtained are evaluated) (Brown, 1987; Cheng & Chan, 2021; Jacobs & Paris, 1987; Schraw & Dennison, 1994).

In the present study, we focus on learners' metacognitive knowledge and regulation as measured by the MAI (Schraw & Dennison, 1994) because it is widely considered the gold standard for measuring the construct (Akın et al., 2007; Harrison & Vallin, 2018). This instrument has been used in almost all existing languages with adequate Cronbach's alpha coefficients and construct validity outcomes. Table 1 shows some of the main background studies on the psychometric properties of the MAI in the case of university students.

The MAI is the most used instrument to evaluate self-report metacognitive knowledge and regulation in different languages due to its easy and quick administration, as well as its collective application (which saves time for the researcher) and low cost. In this context, although it is a tool that was originally created in English (Schraw & Dennison, 1994) and has been translated into almost all languages, the studies referring to the validation and standardization of the MAI in Spanish are scarce and incongruent regarding their results. Table 2 presents research attempts to adapt the instrument to the specifications of the Spanish language.

The present study constitutes the first effort to adapt and standardize the MAI (Schraw & Dennison, 1994) into a standard, international Spanish that can be used with confidence to evaluate undergraduate university students in different cultural contexts in Ibero-America. For this purpose, the researchers employed a process of translation, back translation, and agreement on the most appropriate linguistic form of the different items, so that their use could be relevant in different cultural contexts and Spanish-speaking countries, a process in which the guidelines proposed by the International Test Commission (ITC) were considered to search for psychometric evaluation tasks at an international level (Elosua, 2017; Hambleton, 1996; Hernández et al., 2022; Muñiz et al., 2013, 2015, 2016).

Table 1 Research regarding the psychometric properties of the metacognitive awareness inventory

Authors	Country	Language	Sample Size	Cronbach's α	Version
Akın et al. (2007).	Turkey	Turkish	*N*=607	Full Scale, .95	Final version with 52 items
Lima Filho & Bruni (2015)	Brazil	Portuguese	*N*= 1,058	KoC=.83 RoC=.84	Final version with 52 items
Masoodi (2020)	Lithuania	Farsi	*N* Lithuania=296 Iran=459 *N* total=755	Full Scale= 88 KoC=.85 RoC=.92	Final version with 52 items
Xethakis (2020).	Japan	Japanese	*N*=729	KoC=.84 RoC=.84	Final version with 52 items
Omprakash et al. (2021)	India	Hindi, English	*N*=933	Full Scale, .93	Final version with only 40 of the original 52 items
Perikova & Byzova (2022).	Russia	Russian	*N*=527	KoC=.81 RoC=.82	Final version with only 32 of the original 52 items

Key: KoC Knowledge of Cognition, *RoC* Regulation of Cognition

Table 2 Research regarding the psychometric properties of the metacognitive awareness inventory in spanish-speaking countries

Authors	Country	Participants	Sample Size	Cronbach's α	Version
Huertas et al. (2014)	Colombia	Middel School and High School Students (14-18 years)	*N*=536	Full Scale=.94	Validated final version on the 52 items, rating form on a Likert Scale from: 1. Completely disagree, to 5. Completely agree.
Ulloa Ordaya (2019)	Peru	Middle School and High School Students (11-18 years)	*N*=554	McDonald's Omega (in lieu of α) KoC ω=.886 RoC ω=.934	Validated final version on the 52 items, rating form on a Likert Scale from: 1. Completely disagree, to 5. Completely agree.
Gutierrez de Blume & Montoya Londoño (2021)	Colombia	Undergraduate University Students in Education (20- 30 years)	*N*=528	KoC=.86 RoC=.93 McDonald's Coefficient KoC ω=.83 RoC ω=.90	Validated final version on the 52 items, the score for each item was marked on a vertical bar in the form of a continuous bipolar line from 0-100 (where 0 represents "not at all true for me" to 100 which means "very true for me.")
González-Cabañes et al. (2022)	Spain	Undergraduate University Students (17- 54 years)	*N*=1,076	Study 1 KoC=.81 RoC=.77 Study 2 KoC=.81 RoC=.81	Validated final version on only 19 of the original 52 items, on a Likert scale from 5 as "very typical of me" to 1 = "not at all typical of me").

Key: KoC Knowledge of Cognition, *RoC* Regulation of Cognition

Method

Participants

The study employed a convenience sampling approach in which 1,622 undergraduate students from 12 Latin American cultures and one European culture (Spain) participated. Participants were pursuing an undergraduate degree during the year 2021-2022 and all of them voluntarily agreed to participate in the study. No incentives were provided to participants for their inclusion in the study.

Regarding gender, 699 identified as male and 923 as female. All students met the following inclusion criteria: 1) age between 18 and 25 years ($M = 20.93$; $Median = 20$; $SD = 6.01$); 2) absence of repetition or school lag; and 3) they had to have completed informed consent for their involvement in the research process. Thus, the sociodemographic-variable distribution was typical of each of the participating universities. Table 3 includes additional sociodemographic characteristics of the sample.

Procedure

Process Phase

1) Translation, review, and development stage of the test:

   First, the original English language instrument was consulted, available and freely accessible to the researchers, through the publication of the original research study (Schraw & Dennison, 1994). Next, this original version was reviewed in its final form to begin the translation process by one of the group's researchers, who was trained directly in his doctoral studies by one of the original authors of the MAI, Dr. Gregory Schraw. The original English version we employed for translation is displayed in Appendix A and the Spanish version we translated using the translation-back translation method is found in Appendix B.

   Subsequently, for the translation and back-translation process of the final form of the test in Spanish, a commission was formed from the International Metacognition Group (https://metacog-global.com) made up of 22 experts in research centered on metacognition, 20 with doctoral degrees, and two doctoral students, among whom are psychologists, a graduate in pedagogy, a statistician, and an anthropologist, and some of these researchers with competence in both languages (Spanish and English).

   An extensive process of adapting the items was executed, including reviewing the forms, equivalences, and linguistic correction. The final form of the test approved by the 22 researchers was compared with previous psychometric studies of the instrument already published and available in Spanish for its original form of 52 items (Gutierrez de Blume & Montoya Londoño, 2021; Huertas et al., 2014; Muñiz et al., 2016; Ulloa Ordaya, 2019). However, the present study included a sample of students from a significant number of Spanish-speaking countries: Argentina, Chile, Uruguay, Bolivia, Peru, Panama, Colombia, Mexico, Costa Rica, Guatemala, Puerto Rico, Spain, and Ecuador.

   For the final translated version, the format of the items, the response scale, and the method of administration and scoring of the tool were modified from the original English version (Schraw & Dennison, 1994). Rather than rely on an ordinal Likert scale, we employed a truly continuous scale from 0-100 for each item, instructing participants

Table 3 Sociodemographic variables of the sample

Variable	Frequency	%
Highest Educational Level of the Family		
Grade/Elementary School	138	9.4
High School or Equivalent	357	18.9
Professional Degree	312	15.9
Vocational Degree	151	10.3
Some College	147	10.1
Baccalaureate Degree	300	20.5
Graduate Degree (Masters or Doctorate)	217	14.9
Year at University		
First	484	33.1
Second	374	25.6
Third	357	19.0
Fourth	265	12.6
Fifth	142	9.7
Type of University		
State-subsidized/Public	975	61.1
Private	647	38.9
Academic Program of Study		
Natural Sciences	237	16.2
Social Sciences	863	59.1
Developmental and Construction Sciences	125	8.6
Fine Arts and Humanities	155	5.1
Health Sciences	171	6.2
Other	71	4.9

$N = 1{,}461$

that any value from 0-100 was valid, and we employed two qualitative anchors on each extreme of the continuum: "0 indicates the statement is not at all true of me"; and "100 indicates the statement is completely true of me". Per the advice of Schraw (2009b), using truly continuous (ratio) scales is superior to ordinal Likert scales, and it improves psychometric parameters of measurement, an approach that was successful in previous studies (Gutierrez & Price, 2017; Gutierrez & Schraw, 2015). In addition, a pilot study was conducted, a process in which it was verified that the items were understood by students from the countries that were included in the study, with an approximate time to complete the items between 20 and 30 minutes. This pilot study included 75 university undergraduate students in which a Google Form was distributed with open-ended questions related to clarity of individual items.

2) Confirmation and test application stage:

Prior to full-scale administration, we evaluated the psychometric properties of the proposed version, adjusted to standard, international Spanish. For this process, we relied on the *Standards for Educational and Psychological Testing's* (AERA et al., 2014) guidelines for establishing validity evidence via an exploratory factor analysis, with common factor extraction and an oblique rotation, given that the factors were expected to be highly correlated. In this psychometric analysis of the pilot test, the evaluation of the

items showed adequate anticipated internal structure of the MAI, with Cronbach's alpha coefficients ranging from .794 to .911, showing correlations in the expected theoretical direction (i.e., positive).

In all countries, the application of the test was done by the main researchers of the study, almost all of them with a postgraduate degree or doctorate in psychology, or related areas, who had extensive experience in the application of psychometric tests. The process in which the proposed guidelines for the application of this type of task in educational contexts were also considered (Hambleton, 1996).

3) Stage of establishing scores and interpretation:

Field work was done with the application of the MAI with samples of students from all participating countries. For this purpose, students from different universities where there was a researcher who was participating in the standardization process from their respective country were recruited. Sampling was by convenience and the students were recruited by the researcher responsible for each country, who first held a meeting to explain the objective of the study and administer the informed consent process. Once informed consent was understood and signed by the students who agreed to participate, there was support from teachers from said universities, who authorized the administration of the instrument collectively in their classes during the first and second semester of 2022.

The application of the instrument was administered with the directions and guidelines given by the country's researcher and online, through a questionnaire that was uploaded to the Qualtrics platform (link to the electronic survey), with an approximate duration of half an hour. Participants were permitted to toggle back and forth between pages of the Qualtrics survey and items were offered in blocks to avoid fatigue. The ethical guidelines for conducting research with human participants of each respective country were adopted for informed consent.

Data analysis

Before performing the data analysis process, outlier detection and assumption testing procedures were performed, such as univariate and multivariate normality tests, reproducibility of correlation matrices, and lack of multicollinearity. The data did not contain univariate or multivariate outliers and they met all the required statistical assumptions, except multivariate normality (Mardia's Normalized Estimate [MNE] = 20.636). Therefore, corrections were necessary for this violation of multivariate normality. According to Bentler (2005), any value greater than six for MNE is considered a multivariate value that does not comply with the principles of multivariate normality, and a greater distance from six indicates greater degrees of lack of multivariate normality, which is why robust maximum likelihood (RML) statistics, instead of normally distributed statistics, were used. RML statistics adjust the standard errors of parameter estimates to account for the magnitude of multivariate nonnormality; that is, through this procedure a greater adjustment of the standard errors is made for more serious violations of multivariate normality. The EQS 6.4 software (Bentler, 2005) was used to perform a standard confirmatory factor analysis (CFA).

We randomly split the total sample of 1,622 participants into two separate samples using the random number generator subcommand in SPSS 28. This yielded two randomly generated samples of participants, each sample consisting of 811 cases. First, we employed Sample 1 for initial calibration sample analyses, as we describe next. Subsequently, we

conducted cross-validation analyses with Sample 2 to examine whether the CFA model we evaluated in Sample 1 is valid in a different sample from the same population.

Adjusted RML goodness-of-fit indices (*NNFI, *CFI, *IFI) ≥ .90 suggest an adequately fitting model, and those ≥ .95 suggest excellent fit of the model to observed data. With respect to residuals, standardized root mean square residual (SRMR) values ≤ .08 suggest reasonable errors in estimating model parameters and root mean square error of approximation (*RMSEA) values ≤ .08 suggest that the model parameters approximate those of the population adequately and those < .05 suggest good fit to the data (Byrne, 2006; Kline, 2005). *Dillon-Goldstein's rho* (ρ) was also used to assess the overall or composite reliability of the model. *Rho* measures how well the manifest/indicator variables, as a block, represent the latent variable in which they are hypothesized to load. Like the interpretation of Cronbach's alpha, higher values for *rho* indicate greater model reliability, with .70 serving as the lower-bound for adequate model reliability (Werts et al., 1974).

Results

Baseline standard CFA model for the 12 participating countries with sample 1

Descriptive statistics for the entire sample are presented in Table 4 and Figure 1 (1a-1c) includes sample correlation heatmaps (Bosworth et al., 2017) of the eight subscales of the MAI. For interpretation purposes, we linearly transformed the variables to a 10-point scale for ease of viewing, as heatmaps are best suited to ordinal or nominal scaling. The zero-order correlations between the two factors of knowledge and regulation of cognition ranged between $r = 0.53$ and $r = 0.72$ between countries, indicating the absence of collinearity in the data. As is evident from Table 4, participants from these Spanish-speaking countries generally rated themselves rather highly on their self-report metacognitive skills.

Table 4 Descriptive statistics and cronbach's alpha coefficients of metacognitive variables by scale for the sample

Variables	*M*	*SD*	α
Subjective Metacognitive Awareness			
MAI_KoC	64.11	10.81	0.91
MAI_Dec	60.98	13.58	0.86
MAI_Proc	61.16	15.64	0.77
MAI_Cond	64.33	14.35	0.78
MAI_RoC	62.67	11.66	0.90
MAI_Plan	57.38	13.85	0.79
MAI_Mon	61.88	13.20	0.80
MAI_IM	62.56	12.54	0.77
MAI_Deb	69.01	14.06	0.79
MAI_Eval	61.38	14.52	0.75

$N = 1{,}622$

Key: *M* Mean, *SD* Standard deviation, α Cronbach's alpha coefficients, Cronbach's alpha, MAI Metacognitive Awareness Inventory, *KoC* Knowledge of cognition, *Dec* Declarative knowledge, *Proc* Procedural knowledge, Cond Conditional knowledge, *RoC* Regulation of cognition, *Plan* Planning, *Mon* Comprehension monitoring, *IM* Information management, *Deb* Debugging, *Eval* Evaluation

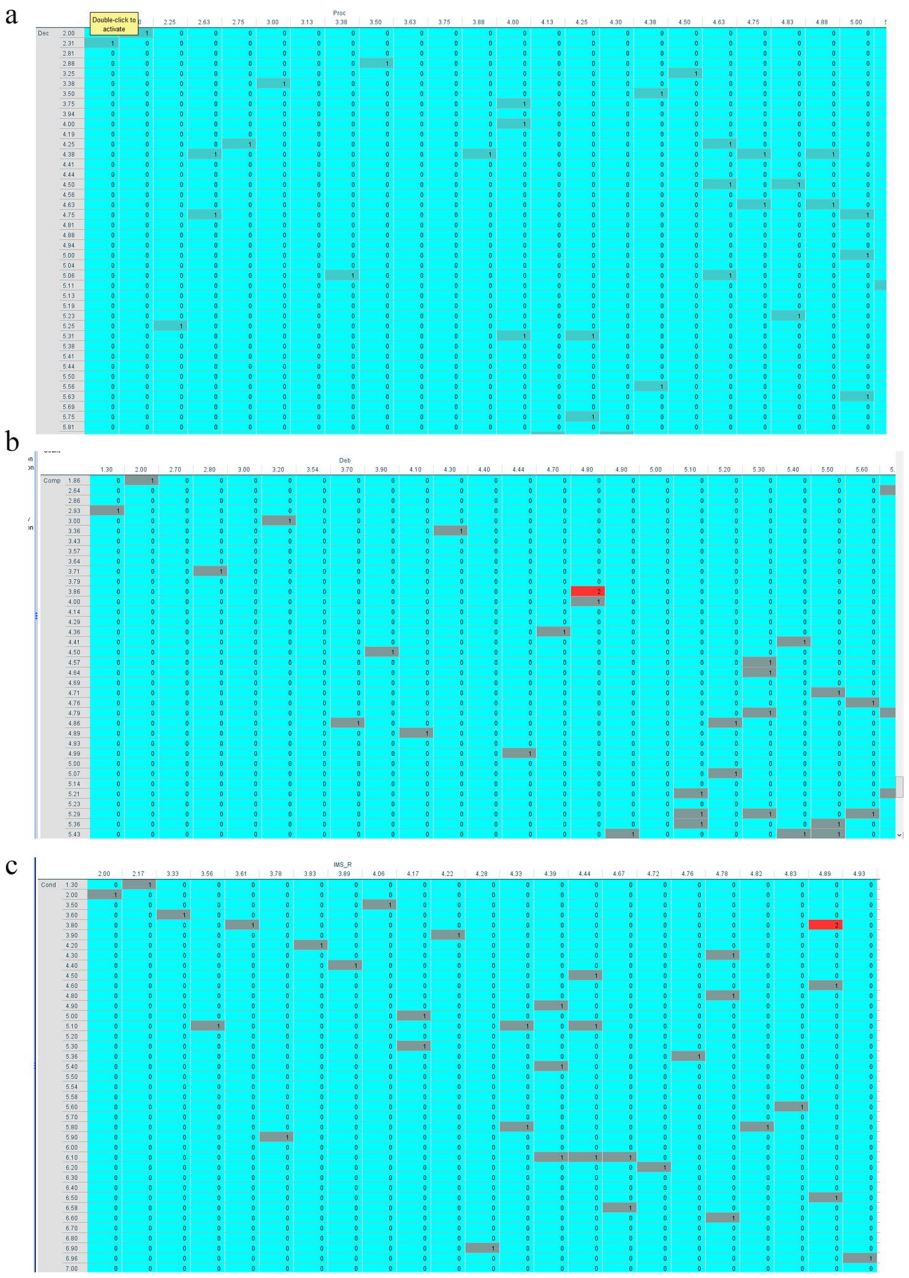


Figure 1 Sample correlation heatmaps of some of the subscales of the MAI. *Note Figure 1a*. Correlation heatmap between declarative and procedural knowledge within the knowledge of cognition scale. *Note Figure 1b*. Correlation heatmap between comprehension monitoring and debugging within the regulation of cognition scale. *Note Figure 1c*. Correlation heatmap between conditional knowledge in the knowledge of cognition subscale and information management within the regulation of cognition scale.

On a continuous scale from 0 to 100, in which higher values represent greater self-perceived metacognitive awareness, participants consistently rated their average metacognitive awareness well above the median of 50. Interestingly, participants reported the highest and lowest mean scores on two regulation of cognition subscales, with debugging strategy use as the highest mean and planning as the lowest mean.

The initial CFA model specified two factors: knowledge of cognition and regulation of cognition, as originally proposed by Schraw and Dennison (1994) in the English validation of the MAI. The results of the initial CFA solution showed a model that did not indicate the best fit to the observed data, S-B χ^2 ($N = 811$; $df = 1{,}273$) = 3952.04, $p < .001$, *NNFI = .813, *CFI = .801, *IFI = .800, SRMR =.056, RMSEA = .059 [$CI_{90\%}$ = .051, .061]. Standardized factor loadings for this initial model ranged from 0.386 ($R^2 = 0.149$) to 0.611 ($R^2 = 0.373$), with the correlation between the two factors being moderate-to-high, $r = 0.603$. Inspection of the Lagrange Multiplier (LM) test for adding parameters suggested that adding several error covariances to the model would significantly improve its fit. As items from each respective subscale are theoretically expected to share significant residual variance (Schraw, 2009b), we added two error covariances from the planning subscale (MAI22: "I ask myself questions about the material before I begin."; MAI42: "I read instructions carefully before I begin a task.") and two from the evaluation subscale (MAI19: "I ask myself if there was an easier way to do things after I finish a task."; MAI24: "I summarize what I've learned after I finish."). Because these additional error correlations made substantive theoretical sense, these parameters were added to a respecified CFA model.

The respecified base model for Sample 1, with all theoretically-grounded error covariances added to the model, fit the data exceptionally well, S-B χ^2 ($N = 811$; $df = 1{,}220$) = 2573.12, $p < .001$, *NNFI = .964, *CFI = .968, *IFI = .965, SRMR =.047, RMSEA = .040 [$CI_{90\%}$ = .036, .043]. The standardized factor loadings for each of the 52 manifest variables for this base model were statistically significant and ranged from 0.493 ($R^2 = 0.243$) to 0.741 ($R^2 = 0.549$). The factors of knowledge of cognition (KoC) and regulation of cognition (RoC) were positively and significantly correlated with each other, $r = 0.643$. The Dillon-Goldstein *rho*, as an indicator of the reliability of the composite model, was 0.951 for the base CFA model, indicating high reliability of the composite model (Werts et al., 1974). Since no other new respecifications made theoretical sense, this was considered the final model. Standardized factor loadings and for this final baseline model are presented in Table 5.

Cross-validation of the final CFA model with sample 2

This final CFA model from Sample 1 was subsequently evaluated for its fit to the randomly generated Sample 2. As with Sample 1, Sample 2 consisted of 811 participants. As with Sample 1, the cross-validated CFA model showed excellent fit to the observed data, S-B χ^2 ($N = 811$; $df = 1{,}220$) = 2880.48, $p < .001$, *NNFI = .948, *CFI = .957, *IFI = .951, SRMR =.044, RMSEA = .037 [$CI_{90\%}$ = .033, .041]. Standardized factor loadings for the cross-validated model were also all statistically significant and ranged from 0.411 ($R^2 = 0.169$) to 0.694 ($R^2 = 0.482$). The two factors, KoC and RoC, were likewise positively and strongly correlated, Pearson's $r = 0.667$, with Dillon-Goldstein *rho* (0.948) indicating high composite model reliability. This converging evidence indicates that the final CFA representing the factor structure of the MAI in these Spanish-speaking samples is valid and can be generalized to other samples from the same population.

Measurement invariance tests across nine countries

The measurement invariance of this final baseline model was then examined for the entire sample across the 12 countries. However, only nine of the 12 countries (Argentina, Bolivia, Colombia, Costa Rica, Guatemala, Panama, Peru, Spain, and Uruguay) recruited a sufficiently large sample size (i.e., $n \geq 100$) to justify a multigroup CFA analysis, as Bentler (2005) and Kline (2005) warn that CFA analyzes with small sample sizes can introduce biased parameter estimation. Therefore, the following measurement invariance analyzes were applied only to these nine countries.

A multigroup CFA was performed to assess the invariance of the structural path coefficients across the nine countries with a sufficiently large sample size. First, a fully constrained and fully saturated baseline model was established for all nine groups to examine the feasibility of the hypothesized CFA model, specifying the direct paths and imposing equality constraints on all coefficients and covariances of structural path coefficients.

Equality constraints were subsequently removed individually for each parameter (i.e., freely estimated) that reached statistical significance at the $p < 0.05$ level, using the LM χ^2 multivariate test to univariate increment to release equality constraints. This procedure was repeated until no univariate increase in LM χ^2 of other parameters reached statistical significance. This model was then considered the final model. Releasing equality constraints for any given parameter indicates that the parameter in question differs statistically significantly between two or more countries. Finally, the $\Delta\chi^2$ (chi-square difference) test was performed to compare the null model (i.e., fully constrained, fully saturated) and the final model (i.e., equality constraints released).

The reference model (same specification as the final model for the entire sample presented above) used for the nine countries fit the observed data reasonably well, S-B χ^2 ($N = 1{,}401$; $df = 1{,}220$) = 2563.26, $p < .001$, *NNFI = .946, *CFI = .949, *IFI = .951, SRMR =.041, RMSEA = .035 [$CI_{90\%}$ = .032, .037]. The standardized factor loadings for each of the 52 manifest variables for this base model were statistically significant and ranged from 0.496 ($R^2 = 0.246$) to 0.803 ($R^2 = 0.645$). The factors of knowledge of cognition and regulation of cognition were positively and significantly correlated with each other, $r = 0.644$. The Dillon-Goldstein *rho*, as an indicator of the reliability of the composite model, was 0.949 for the base CFA model, indicating high reliability of the composite model.

The results of measurement invariance tests revealed that the correlation between knowledge of cognition and regulation of cognition factors varied in some countries compared to the baseline model described above. The only other parameters that varied between some countries were two manifest variables on the planning subscale and two manifest variables on the comprehension monitoring subscale. The results of these parameter variations are presented in Table 6.

Despite these variations across the nine Spanish-speaking countries, the results of the $\Delta\chi^2$ test between the fully constrained baseline model and the final model with five equality constraints released were not statistically significant, $p = 0.12$. Therefore, the factor structure of the MAI using a standard, international Spanish remained invariant between these countries.

Table 5 Standardized factor loadings and explained variance for all 52 items of the final baseline CFA model

Item	RoC SFL	KoC	R^2
MAI1	.74		.55
MAI2	.71		.50
MAI4	.62		.38
MAI6	.60		.36
MAI7	.68		.72
MAI8	.58		.34
MAI9	.73		.53
MAI11	.52		.27
MAI13	.72		.52
MAI19	.49		.24
MAI21	.57		.32
MAI22	.49		.24
MAI23	.64		.41
MAI24	.55		.30
MAI25	.55		.30
MAI28	.71		.50
MAI30	.51		.26
MAI31	.50		.25
MAI34	.66		.44
MAI36	.58		.33
MAI37	.70		.49
MAI38	.50		.25
MAI39	.62		.38
MAI40	.49		.24
MAI41	.54		.29
MAI42	.50		.25
MAI43	.51		.26
MAI44	.57		.32
MAI45	.74		.55
MAI47	.66		.44
MAI48	.69		.47
MAI49	.53		.28
MAI50	.58		.33
MAI51	.57		.32
MAI52	.49		.24
MAI3	.55		.30
MAI5	.61		.37
MAI10		.65	.43
MAI12		.59	.35
MAI14		.68	.47
MAI15		.52	.27
MAI16		.73	.53
MAI17		.49	.24
MAI18		.54	.29
MAI20		.70	.49

Table 5 (continued)

Item	RoC SFL	KoC	R^2
MAI26		.58	.34
MAI27		.56	.32
MAI29		.62	.38
MAI32		.56	.31
MAI33		.63	.40
MAI35		.55	.30
MAI46		.71	.50

Key: KoC = Knowledge of Cognition Factor; RoC = Regulation of Cognition Factor; SFL = Standardized Factor Loading; R^2 = Squared Multiple Correlation Coefficient.

Discussion

The present study supported the two-factor structure of the MAI originally proposed by Schraw and Dennison (1994) in their English-speaking samples from the United States. This is noteworthy because ours is the first study to employ such many Spanish-speaking countries across Latin America and Spain using a standardized, international Spanish that is free from cultural and linguistic nuances. This result is consistent with the findings reported in previous studies, in which the viability of the two major factors of metacognition is confirmed: knowledge and regulation, in research conducted in different geographical and cultural contexts in which the MAI was used as an instrument (Akın et al., 2007; Craig et al., 2020; Harrison & Vallin, 2018; Magno, 2010; Pour & Ghanizadeh, 2017; Teo & Lee, 2012). Having a calibrated, validated instrument, with adequate psychometric properties across so many countries and cultures that communicate in Spanish, that allows researchers and the person evaluated to recognize the levels of knowledge and regulation is useful as a baseline for the evaluation of the state of metacognition.

Likewise, the two-factor model was adjusted for the samples of the participating countries, with minor variations in four items of the planning and monitoring subscales, in the case of Argentina and Spain. However, these variations did not lead to an appreciable difference between the null model and the respecified model with the freely estimated parameters, as the difference between the two models was not statistically significant. This allows us to underscore the usefulness of the version of the MAI proposed in the present study for the evaluation of metacognition in standard, international Spanish. This is possible, to the extent that the unity of Spanish, within its diversity, is currently recognized with regard to the set of linguistic levels, including, at the lexical level, as common to a large part of Spanish speakers (Garrido, 2010; López Morales, 2010; Torres, 2013).

Different researchers have argued that evaluating metacognition is not easy, for many reasons, among which are the diversity of theoretical models and multi-components that vary, typical of the construct, and that make it blurry and diffuse (Azevedo & Aleven, 2013; Buratti & Allwood, 2015; Efklides & Misailidi, 2010; Tarricone, 2011; Tobias & Everson, 2009). This leads the researcher to face different challenges when trying to evaluate self-report metacognition (Ozturk, 2017; Schraw, 2000; Veenman, 2005).

Among the greatest difficulties that the researcher may face when evaluating metacognition are the need to guarantee that people understand the questions asked of them in self-report instruments, that they are willing and open to asking for clarification when required,

or also, to sharing the genuine subjective experiences they are feeling regarding learning or evaluation tasks. In this sense, people's responses to their metacognitive performance could be more a reflection of the insecurity they feel in the moment, due to their lack of self-knowledge regarding their own learning processes, or their desire to communicate a socially desirable tendency in their responses. Although the person does indeed have good performance on metacognitive skills and reports it in the measurement instrument, it is not certain that they can have these skills available to use them when faced with a cognitive or learning task that require self-regulated performance (Baker & Cerro, 2000; Craig et al., 2020; Schraw, 2000; Whitebread et al., 2009).

However, despite the difficulties noted in measuring metacognition, the present study conducted analyses with such a large level of scope, in which people from so many Latin American countries participated, is relevant and important. Further, the confirmation of the factorial structure of the construct, and the adequate psychometric properties of the MAI, provide a calibrated measure for the subjective evaluation of metacognition, which is expected to be used with confidence by researchers, teachers, and students of Spanish-speaking countries. As the MAI is a simple tool, which can be applied offline, in person, or remotely, and which has several benefits for its effective use, such as agile completion, economy in time, low cost, and accessibility to all types of populations, even rural or distant, in terms of geographic access, or opportunities for technological access to the Internet (Ozturk, 2017; Veenman, 2005).

Table 6 Variation in parameter estimates between the constrained baseline model of the nine countries and the model in which they were freely estimated

Country	Model	
	Constrained Baseline	Freely Estimated
Correlation between KoC and RoC Factors		
Bolivia	.644	.598
Colombia		.698
Perú		.700
Spain		.552
MAI Item 8*: I set specific goals before starting a task.		
Argentina	.612	.563
Spain		.519
MAI 42*: I read instructions carefully before starting a task.		
Bolivia	.433	.520
Costa Rica		.509
Spain		.411
MAI 1+: I constantly ask myself if I am meeting my goals.		
Argentina	.430	.406
MAI 11+: I wonder if I have considered all the options when I have to solve a problem.		
Argentina	.527	.477
Spain		.481

Note. All parameter estimates for the manifest variables of the MAI represent standardized factor loadings.

Key: KoC Knowledge of Cognition, *RoC* Regulation of Cognition, * Planning Subscale, + Comprehension Monitoring Subscale

It is important to note as desirable the need to increasingly have metacognition evaluation protocols that involve well-calibrated subjective measures and objective performance measures. For this, the best available resource is the MAI, as in almost all the cultural contexts in which it has been used, researchers from many places in the world have provided evidence to support the factorial structure. These studies have reported appropriate Cronbach's alpha coefficients and construct validity.

Methodological reflections, limitations, and future research

As with any research involving human participants, there are limitations that merit mention. First, this validation study is based on a self-report measure. This presumes that individuals are not only honest with their responses, but also the best raters of their own metacognitive awareness, neither of which we can guarantee. Second, only nine of the 12 countries collected a sufficiently large sample size (i.e., $n \geq 100$) to conduct the multigroup measurement invariance, making it difficult to generalize findings to those countries. Another challenge with the use of self-report instruments is that they are not necessarily the best measures of actual behavior. Thus, it may be that the metacognitive awareness captured by the MAI may not actually reflect participants' actual metacognitive abilities.

Despite these limitations, however, we would like to highlight the strengths of the present study. First, this is only the second study (the other we found was Gutierrez de Blume et al., 2023), to our knowledge, that has used such a robust sample size across so many countries, enhancing the utility value of the findings from the present study. Thus, we think it is a worthy contribution to research on the measurement of metacognitive awareness.

What remains pending for research groups that work in metacognition in the world is to continue working to calibrate objective measures of metacognition based on performance. This is much needed research to decrease the reliance of researchers on self-report measures that are often fraught with too high a measurement error. Further, objective measures would assist in better aligning measurement with theory. In addition, as the MAI has never been employed as a diagnostic assessment for individual-level inferences, future research should explore potential for individual-level bias in the MAI via differential item functioning, which would permit for the calculation of individual scores. Finally, it is worth noting that future research should examine the similarities between the item intercepts and slopes between the original English version of the instrument and the Spanish one we employed for the present study. Such information would enable researchers to know whether the two versions are actually tapping into the same construct, and hence, interchangeable (i.e., direct comparisons can be made between populations).

Appendix A

Original English Version of the MAI

Metacognitive Awareness Inventory

Directions: Please move the slider to the point on the continuous line under each statement that best corresponds to how true each statement is about you.

Not at all true of me — **Completely true of me**

0———————————————————100

For instance, the closer the slider is to "Not at all true of me" the LESS true that statement is about you. Conversely, the closer the slider is to "Completely true of me" the MORE true that statement is about you. Likewise, moving the slider to either end of the line (0 or 100) indicates that the statement is either not at all true of you (0) or completely true of you (100).

1. I ask myself periodically if I am meeting my goals.

Not at all true of me — **Completely true of me**

0———————————————————100

2. I consider several alternatives to a problem before I answer.

Not at all true of me — **Completely true of me**

0———————————————————100

3. I try to use strategies that have worked in the past.

Not at all true of me — **Completely true of me**

0———————————————————100

4. I pace myself while learning in order to have enough time.

Not at all true of me — **Completely true of me**

0———————————————————————————100

5. I understand my intellectual strengths and weaknesses.

Not at all true of me **Completely true of me**

0———————————————————————————100

6. I think about what I really need to learn before I begin a task.

Not at all true of me **Completely true of me**

0———————————————————————————100

7. I know how well I did once I finish a test.

Not at all true of me **Completely true of me**

0———————————————————————————100

8. I set specific goals before I begin a task.

Not at all true of me **Completely true of me**

0———————————————————————————100

9. I slow down when I encounter important information.

Not at all true of me **Completely true of me**

0———————————————————————————100

10. I know what kind of information is most important to learn.

Not at all true of me **Completely true of me**

0———————————————————————————100

11. I ask myself if I have considered all options when solving a problem.

12. I am good at organizing information.

13. I consciously focus my attention on important information.

14. I have a specific purpose for each strategy I use.

15. I learn best when I know something about the topic.

16. I know what the teacher expects me to learn.

17. I am good at remembering information.

0———————————————————————————100

18. I use different learning strategies depending on the situation.

Not at all true of me — **Completely true of me**

0———————————————————————————100

19. I ask myself if there was an easier way to do things after I finish a task.

Not at all true of me — **Completely true of me**

0———————————————————————————100

20. I have control over how well I learn.

Not at all true of me — **Completely true of me**

0———————————————————————————100

21. I periodically review to help me understand important relationships.

Not at all true of me — **Completely true of me**

0———————————————————————————100

22. I ask myself questions about the material before I begin.

Not at all true of me — **Completely true of me**

0———————————————————————————100

23. I think of several ways to solve a problem and choose the best one.

Not at all true of me — **Completely true of me**

0———————————————————————————100

24. I summarize what I've learned after I finish.

25. I ask others for help when I don't understand something.

26. I can motivate myself to learn when I need to.

27. I am aware of what strategies I use when I study.

28. I find myself analyzing the usefulness of strategies while I study.

29. I use my intellectual strengths to compensate for my weaknesses.

30. I focus on the meaning and significance of new information.

0———————————————————————————100

31. I create my own examples to make information more meaningful.

Not at all true of me — **Completely true of me**

0———————————————————————————100

32. I am a good judge of how well I understand something.

Not at all true of me — **Completely true of me**

0———————————————————————————100

33. I find myself using helpful learning strategies automatically.

Not at all true of me — **Completely true of me**

0———————————————————————————100

34. I find myself pausing regularly to check my comprehension.

Not at all true of me — **Completely true of me**

0———————————————————————————100

35. I know when each strategy I use will be most effective.

Not at all true of me — **Completely true of me**

0———————————————————————————100

36. I ask myself how well I accomplished my goals once I am finished.

Not at all true of me — **Completely true of me**

0———————————————————————————100

37. I draw pictures or diagrams to help me understand while learning.

Not at all true of me — **Completely true of me**

0———————————————————————————100

38. I ask myself if I have considered all options after I solve a problem.

Not at all true of me — **Completely true of me**

0———————————————————————————100

39. I try to translate new information into my own words.

Not at all true of me — **Completely true of me**

0———————————————————————————100

40. I change strategies when I fail to understand.

Not at all true of me — **Completely true of me**

0———————————————————————————100

41. I use the organizational structure of the text to help me learn.

Not at all true of me — **Completely true of me**

0———————————————————————————100

42. I read instructions carefully before I begin a task.

Not at all true of me — **Completely true of me**

0———————————————————————————100

43. I ask myself if what I am reading is related to what I already know.

Not at all true of me — **Completely true of me**

0———————————————————————————100

44. I reevaluate my assumptions when I get confused.

Not at all true of me — **Completely true of me**

0 ——— 100

45. I organize my time to best accomplish my goals.

Not at all true of me — **Completely true of me**

0 ——— 100

46. I learn more when I am interested in the topic.

Not at all true of me — **Completely true of me**

0 ——— 100

47. I try to break studying down into smaller steps.

Not at all true of me — **Completely true of me**

0 ——— 100

48. I focus on overall meaning rather than specifics.

Not at all true of me — **Completely true of me**

0 ——— 100

49. I ask myself questions about how well I am doing while I am learning something new.

Not at all true of me — **Completely true of me**

0 ——— 100

50. I ask myself if I learned as much as I could have once I finish a task.

Not at all true of me — **Completely true of me**

0——————————————————————————100

51. I stop and go back over new information that is not clear.

Not at all true of me — **Completely true of me**

0——————————————————————————100

52. I stop and reread when I get confused.

Not at all true of me — **Completely true of me**

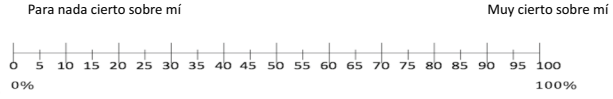
0——————————————————————————100

Appendix B

Spanish Version of the MAI Translated via the Translation-Back Translation Method

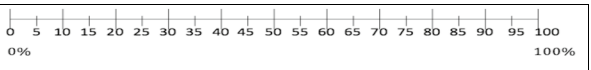
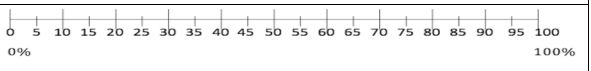
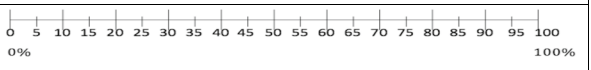
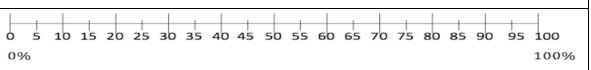
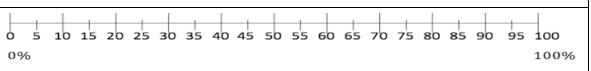
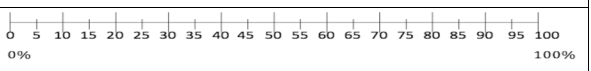
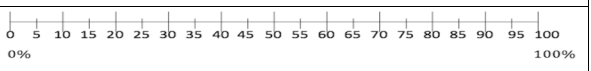
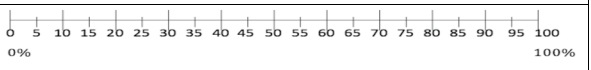
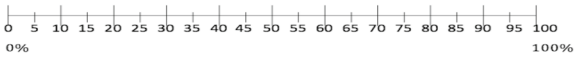
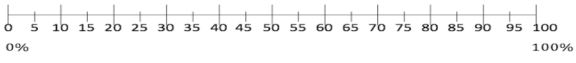
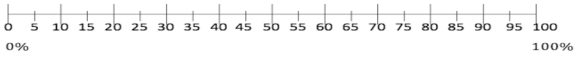
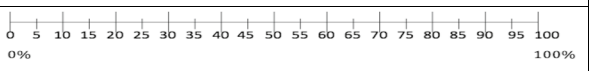
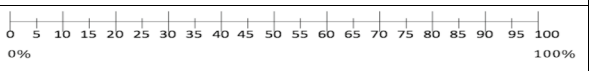
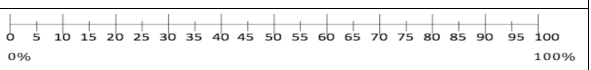
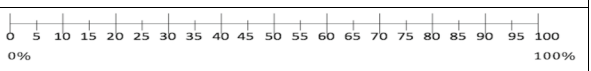
Inventario de Conocimiento Metacognitivo (Schraw & Dennison, 1994)

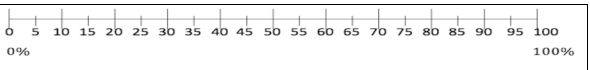
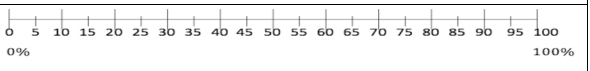
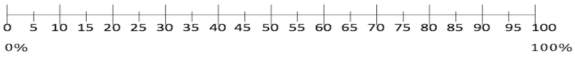
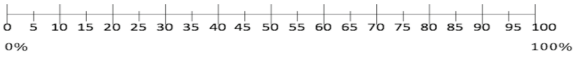
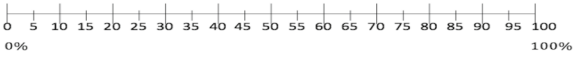
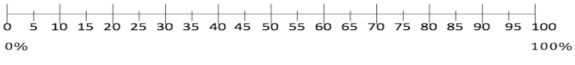
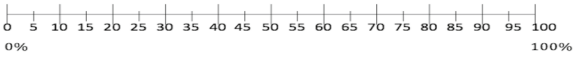
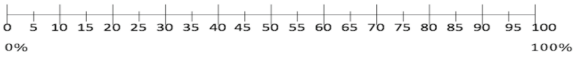
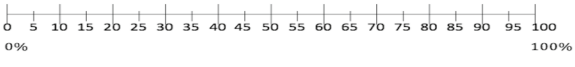
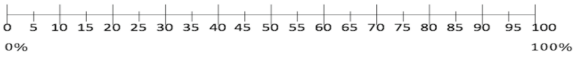
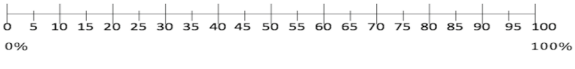
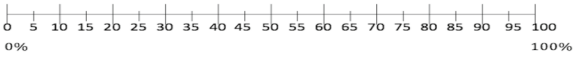
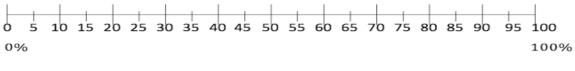
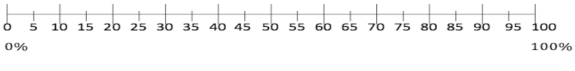
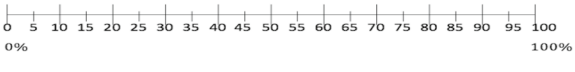
Indicaciones: Por favor dibuje una línea vertical sobre la línea continua en cada uno de los enunciados, que mejor corresponda a qué tan cierto es cada una de las siguientes indicaciones sobre usted.

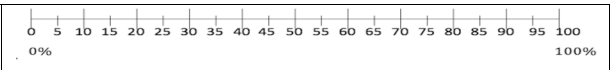
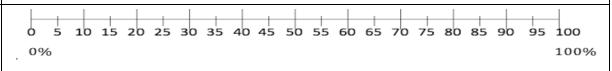
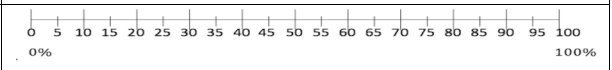
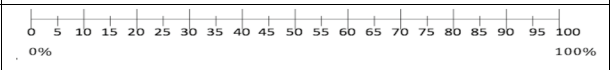
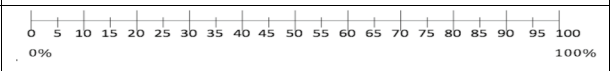
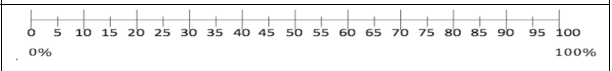
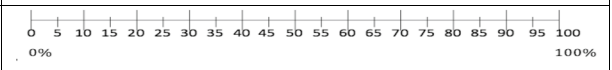
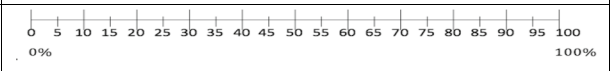
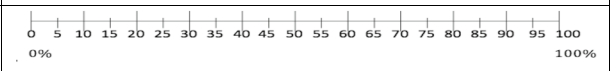
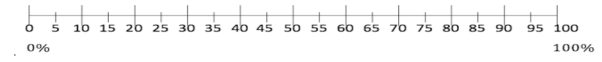
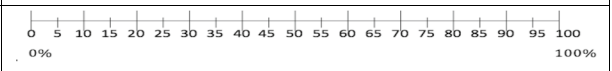
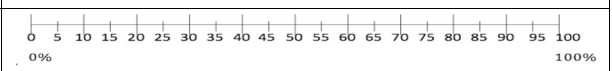
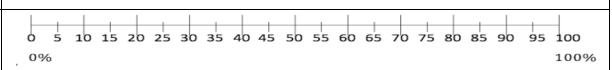
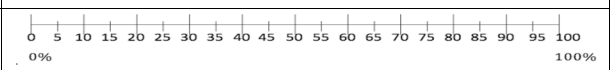


Por ejemplo: Entre más cerca este la línea de "Para nada cierto sobre mí", menos cierto es el enunciado sobre usted, por el contrario, entre más cerca este la línea a "muy cierto sobre mí", más verdadero será el enunciado sobre usted. De la misma manera, dibujar la línea muy al comienzo o muy al final 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.	
10. Sé cuál es la información más importante que debo aprender.	
11. Me pregunto si he tenido en cuenta todas las opciones, cuando tengo que resolver un problema.	
12. Soy bueno organizando información.	
13. Conscientemente enfoco mi atención, en la información importante.	
14. Tengo un propósito específico con cada una de las estrategias que utilizo.	
15. Aprendo mejor cuando se algo con respecto al tema.	
16. Se lo que el profesor espera que yo aprenda.	
17. Soy bueno recordando información.	
18. Utilizo diferentes estrategias de aprendizaje dependiendo de la situación.	
19. Después de terminar una tarea, me pregunto si había una forma más fácil de resolverla.	
20. Tengo control sobre qué tan bien aprendo.	
21. Periódicamente estoy estudiando para ayudarme a comprender relaciones importantes.	
22. Me hago preguntas acerca de las lecturas, antes de comenzar a leer.	
23. Pienso en varias formas de resolver un problema y selecciono la mejor.	

24. Hago resúmenes de lo que he aprendido una vez termino.	
25. Pido la ayuda de otros cuando no comprendo algo.	
26. Puedo motivarme a aprender lo que necesito aprender.	
27. Soy consciente de que estrategias debo utilizar cuando estudio.	
28. Puedo analizar la utilidad de las estrategias que uso cuando estudio.	
29. Utilizo mis fortalezas intelectuales, para compensar mis debilidades.	
30. Me enfoco en el significado y significancia de la información novedosa.	
31. Puedo generar mis propios ejemplos, para que la información sea más significativa.	
32. Puedo juzgar muy bien, que tan bien comprendo una temática o tema.	
33. Por lo general, utilizo estrategias de aprendizaje útiles automáticamente.	
34. Por lo general, puedo disminuir mi ritmo de trabajo para saber si estoy comprendiendo.	
35. Sé cuándo las estrategias que utilizo serán más efectivas.	
36. Puedo saber que tan bien he logrado mis metas, una vez he terminado.	
37. Realizo mapas conceptuales, para ayudarme a comprender mientras estudio.	
38. Me pregunto si he tenido en cuenta todas las opciones, antes de resolver un problema.	

39. Trato de poner toda la información en mis propias palabras.	
40. Cambio las estrategias cuando no logro comprender muy bien.	
41. Utilizo la estructura organizacional del texto para comprender mejor.	
42. Leo las instrucciones cuidadosamente antes de comenzar una tarea.	
43. Me pregunto si lo que estoy leyendo está relacionado con lo que ya se.	
44. Reevalúo lo que he aprendido cuando me confundo.	
45. Organizo mi tiempo para lograr todas mis metas.	
46. Aprendo más cuando estoy interesado en el tema.	
47. Intento estudiar por partes para tener una mejor comprensión.	
48. Me enfoco en los significados generales, más que en los específicos.	
49. Me hago preguntas con respecto a que tan bien estoy haciendo las cosas, cuando aprendo nueva información.	
50. Me pregunto si aprendí tanto como debería, una vez termino la tarea.	
51. Me detengo y puedo volver a revisar información que aún no me es clara.	
52. Me detengo y vuelvo a leer cuando estoy confundido.	

Funding There is no funding to report.

Data availability Data associated with this research are available to anyone upon reasonable request.

Code availability Not applicable.

Declarations

Conflicts of interest/competing interests None of the authors have any conflicts of interest to report.

References

Akın, A., Abacı, R., & Çetin, B. (2007). The validity and reliability study of the turkish version of the metacognitive awareness inventory. *Educational Sciences: Theory and Practice, 7*(2), 671–678. https://toad.halileksi.net/wp-content/uploads/2022/07/metacognitive-awareness-inventory-toad.pdf.

American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (2014). *Standards for educational and psychological testing* (2014 ed.). American Educational Research Association. https://www.aera.net/Publications/Books/Standards-for-Educational-Psychological-Testing-2014-Edition

Azevedo, R., & Aleven, V. (2013). *International handbook of metacognition and learning technologies*. Springer. https://doi.org/10.1007/978-1-4419-5546-3

Baker, L., & Cerro, L. (2000). Assessing metacognition in children and adults. In G. Schraw & J. C. Impara (Eds.), *Issues in the measurement of metacognition* (pp. 99–145). Buros Institute of Mental Measurements.

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology, 52*, 1–26. https://doi.org/10.1146/annurev.psych.52.1.1

Bentler, P. M. (2005). EQS 6: Structural equations program manual. Multivariate Software.

Bosworth, A., Dowall, S., García-Dorival, I., Rickett, N., Bruce, C., Matthews, D., Fang, Y., Aljabr, W., Kenny, J., Nelson, C., Laws, T., Williamson, E., Stewart, J., Carroll, M., Hewson, R., & Hiscox, J. (2017). A comparison of host gene expression signatures associated with infection in vitro by the Makona and Ecran (Mayinga) variants of Ebola virus. *Scientific Reports, 7*(1), 43144. https://doi.org/10.1038/srep43144

Brown, A. (1987). Metacognition, executive control, self-regulation, and other more mysterious mechanisms. In F. Weinert & R. Kluwe (Eds.), *Metacognition, motivación and understanding* (pp. 65–116). Lawrence erlbaum.

Buratti, S., & Allwood, C. (2015). Regulating metacognitive processes-support for a meta-metacognitive ability. In A. Peña-Ayala (Ed.), *Metacognition: Fundaments, applications and trends. A prolife of the current state -of-the-art* (pp. 17–35). Springer.

Byrne, B. M. (2006). Structural equation modeling with EQS: Basic concepts, applications, and programming (2nd ed.). Lawrence Erlbaum Associates Publishers.

Cheng, E., & Chan, J. (2021). *Developing metacognitive teaching strategies through lesson study*. Springer.

Craig, K., Hale, D., Grainger, C., & Stewart, M. (2020). Evaluating metacognitive self-reports: systematic reviews of the value of self-report in metacognitive research. *Metacognition and Learning, 15*, 155–213. https://doi.org/10.1007/s11409-020-09222-y

Cross, D., & Paris, S. (1988). Developmental and instructional analyses of children's metacognition and reading comprehension. *Journal of Educational Psychology, 80*(2), 131–142. https://doi.org/10.1037/0022-0663.80.2.131

Dinsmore, D. L., Alexander, P. A., & Loughlin, S. M. (2008). Focusing the conceptual lens on metacognition, self-regulation, and self-regulated learning. *Educational Psychology Review, 20*, 391–409. https://doi.org/10.1007/s10648-008-9083-6

Double, K., & Birney, D. (2019). Reactivity to measures of metacognition. *Frontiers in Psychology, 10*(2755), 1–10. https://doi.org/10.3389/fpsyg.2019.02755

Dunlosky, J., & Rawson, K. A. (Eds.). (2019). *The Cambridge Handbook of Cognition and Education*. London: Cambridge University Press. https://doi.org/10.1017/9781108235631

Efklides, A. (2006). Metacognition and affect: What can metacognitive experiences tell us about the learning process? *Educational Research Review, 1*(1), 3–14. https://doi.org/10.1016/j.edurev.2005.11.001

Efklides, A. (2008). Metacognition: Defining its facets and levels of func- tioning in relation to self-regulation and co-regulation. *European Psychologist, 13*(4), 277–287. https://doi.org/10.1027/1016-9040.13.4.277

Efklides, A., & Misailidi, P. (2010). Introduction: The present and future in metacognition. In A. Efklides & P. Misailidi (Eds.), *Trends and prospects in metacognition research* (pp. 1–20). Springer.

Elosua, P. (2017). Avances, proyectos y retos internacionales ligados al uso de tests en Psicología. *Estudos de Psicologia Campinas, 34*(2), 201–210. https://doi.org/10.1590/1982-02752017000200002

Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. B. Resnick (Ed.), *The nature of intelligence* (pp. 231–235). Lawrence Erlbaum.

Flores, J., Ostrosky, F., & Lozano, A. (2008). Batería de funciones frontales y ejecutivas: presentación. *Revista de Neuropsicología, Neuropsiquiatría y Neurociencias*, *8*(1), 141–158. Recuperado de https://aalfredoardila.files.wordpress.com/2013/07/ardila-a-ed-2008-funciones-ejecutivas-neuropsicologia-neuropsiquiatria-y-neurociencias-vol-8-n1.pdf. Accessed 19 Mar 2024

Flores-Lázaro, J., Castillo-Preciado, R., & Jiménez-Miramonte, N. (2014a). Desarrollo de funciones ejecutivas, de la niñez a la juventud. *Anales de Psicología, 30*(2), 463–473. https://doi.org/10.6018/analesps.30.2.155471

Flores-Lázaro, J., Ostrosky-Schejet, F., & Lozano-Gutiérrez, A. (2014b). *BANFE - 2. Batería neuropsicológica de funciones ejecutivas y lóbulos frontales*. Manual Moderno.

Follmer, D. J., & Sperling, R. (2016). The mediating role of metacognition in the relationship between executive function and self-regulated learning. *British Journal of Educational Psychology, 86*(4), 1–17. https://doi.org/10.1111/bjep.12123

Garner, R. (1990). When children and adults do not use learning strategies: Toward a theory of settings. *Review of Educational Research, 60*(4), 517–529. https://doi.org/10.3102/00346543060004517

Garrido, J. (2010). Lengua y globalización: inglés global y español pluricéntrico. *Historia y Comunicación Social, 15*, 67–102. https://revistas.ucm.es/index.php/HICS/article/view/HICS1010110063A

González-Cabañes, E., García, T., Álvarez-García, D., Fernández, E., & Rodríguez, C. (2022). Validation of the shortened version of the metacognitive awareness inventory in spanish university students. *Psicothema, 34*(3), 454–462. https://doi.org/10.7334/psicothema2022.75

Gutierrez de Blume, A. P., & Montoya Londoño, D. (2021). Validación y examen de la estructura factorial del *Metacognitive Awareness Inventory* (MAI) en español con una muestra colombiana de estudiantes universitarios. *Revista psicogente, 24*(46), 1–28. https://doi.org/10.17081/psico.24.46.4881

Gutierrez de Blume, A., Schraw, G., Kuch, F., & Richmond, A. (2020). General accuracy and general error factors in metacognitive monitoring and the role of time on task in predicting metacognitive judgments. *Rev CES Psico, 14*(2), 179–208. https://doi.org/10.21615/cesp.5494

Gutierrez de Blume, A. P., Montoya Londoño, D. M., Daset, L., Cuadro, A., Molina Delgado, M., Morán Núñez, O., García de la Cadena, C., Beltrán Navarro, M. B., Arias Trejo, N., Ramirez Balmaceda, A., Jiménez Rodríguez, V., Puente Ferreras, A., Urquijo, S., Arias, W. L., Rivera, L. I., Schulmeyer, M., & Rivera-Sanchez, J. (2023). Normative data and standardization of an international protocol for the evaluation of metacognition in Spanish-speaking university students: A cross-cultural analysis. *Metacognition and Learning, 18*(2), 495–526. https://doi.org/10.1007/s11409-023-09338-x

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

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

Hambleton, R. K. (1996). Adaptación de tests para su uso en diferentes idiomas y culturas: fuentes de error, posibles soluciones y directrices prácticas. En J. Muñiz (Coor.), Psicometría. Universitas.

Harrison, G., & Vallin, L. (2018). Evaluating the metacognitive awareness inventory using empirical factor-structure evidence. *Metacognition Learning, 13*, 15–38. https://doi.org/10.1007/s11409-017-9176-z

Hernández, A., Elosua, P., Fernández-Hermida, J., & Muñiz, J. (2022). Comisión de Test: Veinticinco años velando por la calidad de los test. *Papeles del psicólogo, 43*(1), 55–62. https://doi.org/10.23923/pap.psicol.2978

Huertas, A., Vesga, G., & Galindo, M. (2014). Validación del instrumento "inventario de habilidades metacognitivas (MAI)" con estudiantes Colombianos. *Paxis, 5*, 55–74.

Jacobs, J., & Paris, S. (1987). Children's metacognition about reading: Issues in definition, measurement, and instruction. *Educational Psychologist, 22*(3–4), 255–278. https://doi.org/10.1207/s15326985ep2203&4_4

Kline, R. B. (2005). Principles and practice of structural equation modeling (2nd ed.). Guilford Press.

Lima Filho, R., & Bruni, A. (2015). Metacognitive awareness inventory: Trans- lation and validation from a confirmatory analysis. *Psicologia: Ciência e Profissão, 35*(4), 1275–1293. https://doi.org/10.1590/1982-3703002292013

Livingston, J. A. (2003). *Metacognition: An overview*. (ED474273) ERIC. https://eric.ed.gov/?id=ED474273

López Morales, H. (2010). *La andadura del español por el mundo*. Taurus.

Magno, C. (2010). The role of metacognitive skills in developing critical thinking. *Metacognition and Learning, 5*(2), 137–156. https://doi.org/10.1007/s11409-010-9054-4

Massodi, M. (2020). *Metacognitive awareness in university studies: the comparative study of lithuanian and iranian cases*. Mykolas Romeris University. https://repository.mruni.eu/handle/007/16501. Accessed 19 Mar 2024

Metcalfe, J., & Greene, M. (2007). Metacognition of agency. *Journal of experimental psychology General, 136*(2), 184–99. https://doi.org/10.1037/0096-3445.136.2.184

Muñiz, J., Elosua, P., & Hambleton, R. (2013). Directrices para la traducción y adapta- ción de los tests: segunda edición. *Psicothema, 25*(2), 151–157. https://doi.org/10.7334/psicothema2013.24

Muñiz, J., Hernández, A., & Ponsoda, V. (2015). Nuevas directrices sobre el uso de los tests: investigación, control de calidad y seguridad. *Papeles del Psicólogo, 36*(3), 161–173. https://repositorio.uam.es/handle/10486/671956.

Muñiz, J., Elosua, P., Padilla, J. L., & Hambleton, R. K. (2016). Test adaptation standards for cross-lingual assessment. In C. S. Wells & M. Faulkner-Bond (Eds.), *Educational measurement: From foundations to future* (pp. 291–304). The Guilford Press.

Omprakash, A., Kumar, A., Kuppusamy, M., Sathiyasekaran, B., Ravinder, T., & Ramaswamy, P. (2021). Validation of Metacognitive Awareness Inventory from a Private Medical University in India. *Journal of education and health promotion, 10*, 324. https://doi.org/10.4103/jehp.jehp_39_21

Ozturk, N. (2017). Assessing metacognition: theory and practices. *International Journal of Assessment Tools in Education, 4*(2), 134–148. https://doi.org/10.21449/ijate.298299

Pandey, K., & Mohan, A. (2023). Unveiling the Subtle Strength of "Metacognition" in the domain of Education. *YMER, 22*(3), 1131–1146. https://www.researchgate.net/publication/371176452_Unveiling_the_Subtle_Strength_of_Metacognition_in_the_domain_of_Education. Accessed 19 Mar 2024

Perikova, E., & Byzova, M. (2022). Factor structure of the russian version of the metacognitive awareness inventory. *Kul'turno-istoricheskaya psikhologiya = Cultural-Historical Psychology, 18*(2), 116–126. https://doi.org/10.17759/chp.2022180213

Pintrich, P. R., Wolters, C. A., & Baxter, G. P. (2000). Assessing metacognition and self- regulated learning. In G. Schraw & J. C. Impara (Eds.), *Assessing metacognition and self-regulated learning* (pp. 43–97). Buros Institute of Mental Measurements.

Pour, A. V., & Ghanizadeh, A. (2017). Validating the Persian version of metacognitive awareness inventory and scrutinizing the role of its components in IELTS academic Reading achievement. *Modern Journal Of Language Teaching Methods, 7*(3), 46–63.

Quinn, J., McEachen, J., Fullan, M., Gardner, M., & Drummy, M. (2021). *Sumergirse en el aprendizaje profundo. Herramientas atractivas*. Morata.

Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional Science, 26*(1), 113–125. https://doi.org/10.1023/A:1003044231033

Schraw, G. (2000). Assessing metacognition: Implications of the Buros symposium. In G. Schraw & J. C. Impara (Eds.), *Issues in the measurement of metacognition* (pp. 297–321). Buros Institute of Mental Measurements.

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

Schraw, G. (2009a). A conceptual analysis of five measures of metacognitive monitoring. *Metacognition and Learning, 4*(1), 33–45. https://doi.org/10.1007/s11409-008-9031-3

Schraw, G. (2009b). 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., & Dennison, R. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19*, 460–475. https://doi.org/10.1006/ceps.1994.1033

Serra, M., & Metcalfe, J. (2009). Effective implementation of metacognition. In D. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Handbook of Metacognition in Education* (pp. 278–298). Erlbaum. https://doi.org/10.4324/9780203876428

Tarricone, P. (2011). *The Taxonomy of Metacognition*. Psychology press.

Teo, T., & Lee, C. B. (2012). Assessing the factorial validity of the metacognitive awareness inventory (MAI) in an Asian country: A confirmatory factor analysis. *International Journal of Educational and Psychological Assessment, 10*(2), 92–103.

Tobias, S., & Everson, H. (2009). The importance of knowing what you know: A knowledge monitoring framework for studying metacognition in education. In D. J. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Handbook of Metacognition in Education* (pp. 107–228). Routledge.

Torres, A. (2013). Del castellano de «un pequeño rincón» al español internacional. *Revista de Estudios Lingüísticos Hispánicos, 3*, 205-224. https://ojs.uv.es/index.php/normas/article/view/4679/4538

Ulloa Ordaya, M. (2019). *Propiedades psicométricas del Inventario de Conciencia Metacognitiva (MAI) en estudiantes de secundaria de una institución educativa, UGEL 07, Lima.* [Título profesional en psicología, Universidad Cesar Vallejo, Lima.Perú]. Repositorio Académico de la Universidad Cesar Vallejo. https://repositorio.ucv.edu.pe/handle/20.500.12692/44663

Veenman, M. V. J. (2005). The assessment of metacognitive skills. In B. Moschner & C. Artelt (Eds.), *Lernstrategien und Metakognition: Implikationenfür Forschung und Praxis* (pp. 75–97). Waxmann.

Veenman, M., Van Hout-Wolters, B., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition and Learning, 1*(1), 3–14. https://doi.org/10.1007/s11409-006-6893-0

Veenman, M., Bavelaar, L., De Wolf, L., & Van Haaren, M. (2014). The on-line assessment of metacognitive skills in a computerized learning environment. *Learning and Individual Differences, 29*, 123–130. https://doi.org/10.1016/j.lindif.2013.01.003

Wang, M. C., Haertel, G. D., & Walberg, H. J. (1990). What influences learning? A content analysis of review literature. *The Journal of Educational Research, 84*(1), 30–43. https://doi.org/10.1080/00220671.1990.10885988

Werts, C. E., Linn, R. L., & Jöreskog, K. G. (1974). Intraclass Reliability Estimates: Testing Structural Assumptions. *Educational and Psychological Measurement, 34*(1), 25–33. https://doi.org/10.1177/001316447403400104

Whitebread, D., Coltman, P., Pasternak, D. P., Sangster, C., Grau, V., Bingham, S., & Demetriou, D. (2009). The development of two observational tools for assessing metacognition and self-regulated learning in young children. *Metacognition and Learning, 4*(1), 63–85. https://doi.org/10.1007/s11409-008-9033-1

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

Winne, P. H., & Perry, N. E. (2005). Measuring self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 531–566). Elsevier Academic Press.

Xethakis, L. (2020). The psychometric properties of the Metacognitive Awareness Inventory in the Japanese EFL context. 熊本大学社会文化研究, *18,* 121–146.

Zohar, A. (2020). *Scaling-up higher order thinking*. Springer.

Zohar, A., & Dori, Y. (2012). *Metacognition in Science education. Trends in current research*. Springer.

Authors and Affiliations

Antonio P. Gutierrez de Blume[1] · Diana Marcela Montoya Londoño[2] · Virginia Jiménez Rodríguez[3] · Olivia Morán Núñez[4] · Ariel Cuadro[5] · Lilián Daset[5] · Mauricio Molina Delgado[6] · Claudia García de la Cadena[7] · María Beatríz Beltrán Navarro[8] · Aníbal Puente Ferreras[3] · Sebastián Urquijo[9] · Walter Lizandro Arias[10]

✉ Antonio P. Gutierrez de Blume
agutierrez@georgiasouthern.edu

[1] Department of Curriculum, Foundations, and Reading, Georgia Southern University, P.O. Box 8144, Statesboro, GA 30460-8144, USA

[2] Departamento de Estudios Educativos, Facultad de Artes y Humanidades, Programa de Psicología, Universidad de Caldas, Universidad de Manizales, Manizales, Colombia

[3] Universidad Complutense de Madrid, Madrid, Spain

[4] Universidad de Panamá, Panama City, Panama

[5] Universidad Católica de Uruguay, Montevideo, Uruguay

[6] Universidad de Costa Rica, San José, Costa Rica

[7] Universidad del Valle de Guatemala, Guatemala City, Guatemala

[8] Universidad de Guadalajara, Guadalajara, Mexico

[9] Universidad Mar del Plata, Mar del Plata, Argentina

[10] Universidad Católica San Pablo, Arequipa, Peru