Electronic Journal of Research in Educational Psychology
ejrep
EDUCATION & PSYCHOLOGY
I+D+i
Ilustre Colegio Oficial de Psicología de Andalucía Oriental

EVAPROMES, an assessment scale for metacognitive processes in writing

Virginia Jiménez-Rodríguez[1, 2], María Alexandra Ulate-Espinoza[3], Jesús María Alvarado-Izquierdo[1] & Aníbal Puente-Ferreras[1]

[1] Universidad Complutense de Madrid.
[2] Universidad Camilo José Cela, Madrid.
[3] Ministerio de Educación Pública de Costa Rica, San José.

Spain & Costa Rica

Correspondence: Virginia Jiménez Rodríguez. Facultad de Trabajo Social. Campus de Somosaguas. 28224. Pozuelo de Alarcón. Madrid. Spain.
E-mail: virginiajimenez@psi.ucm.es; María Alexandra Ulate Espinoza m.ulate.espinoza@mep.go.cr

Abstract

Introduction. Writing is considered a basic learning tool today. Its underlying processes are complex, and observation of schoolchildren's poor writing has led to a consideration of new forms of teaching and learning to write. Skilled writers are aware of what their own learning processes are, how they work, and how to optimize performance and control of their own writing processes.

Method. The EVAPROMES test includes 28 items that represent "writing situations", for which writers must select one of three possible responses. The subject's choice reveals "how writers perceive themselves" and "how writers believe that they would act" when facing possible dilemmas in writing, as well as their level of motivation when writing. Each response alternative expresses a different level of writing awareness. The test has been validated with Costa Rican and Spanish samples; the total sample was 1442 students between the ages of 10 and 15 (915 Costa Ricans and 527 Spanish).

Results. The results indicate that EVAPROMES is an instrument that can collect information on children's and adolescents' level of writing awareness, offering precise measurement of implicit processes, and makes possible a realistic prognosis of their performance in writing and in academics in general.

Conclusion. EVAPROMES identifies students with low and high writing awareness. Based on this identification, it is possible to design specific intervention programs in metacognitive strategies in the area of writing, adapting the educational response to the students assessed.

Keywords: Metacognition Processes, Writing, Planning, Self-Regulation, Assessment, Strategies.

Reception: 08.09.15 Initial acceptance: 09.16.15 Final acceptance: 11.19.15

EVAPROMES, una escala para evaluar los procesos metacognitivos en escritura

Resumen

Introducción. La escritura se considera hoy una herramienta básica de aprendizaje. Los procesos que subyacen a ella son complejos y la detección de la pobreza escritora de los estudiantes en las escuelas hace que se planteen nuevas formas de enseñar y aprender a escribir. Un escritor experto es consciente de cuáles son sus propios procesos de aprendizaje, cómo funcionan, y conoce cómo optimizar su funcionamiento y el control de sus propios procesos escritores.

Método. La prueba EVAPROMES incluye 28 ítems que representan "situaciones de escritura" con tres alternativas de respuesta que se valoran ante las que el escritor debe tomar partido. La alternativa seleccionada permite conocer "cómo los escritores se perciben" y "cómo creen que actuarían" ante posibles dilemas que se les plantean a la hora de escribir, así como su grado de motivación ante el acto escritor. Cada alternativa expresa un grado diferente de conciencia escritora. La prueba ha sido validada en muestra costarricense y española, compuesta por 1442 estudiantes entre 10 y 15 años (915 costarricenses y 527 españoles).

Resultados. Los resultados indican que EVAPROMES es un instrumento que permite recoger información del grado o nivel de conciencia escritora en niños y adolescentes, con una medida precisa de los procesos implícitos, permitiendo un buen pronóstico del rendimento en escritura y académico en general.

Conclusión. EVAPROMES permite identificar al alumnado con baja y alta conciencia escritora. Basándose en esta identificación, se pueden diseñar programas de intervención específicos en estrategias metacognitivas en el área de la escritura adaptando la respuesta educativa a los alumnos evaluados.

Palabras Clave: Procesos Metacognitivos, Escritura, Planificación, Autorregulación, Evaluación, Estrategias.

Recibido: 29.08.15 *Aceptación Inicial: 16.09.15* *Aceptación final: 19.11.15*

Introduction

Among the cognitive activities specific to human beings, writing is undoubtedly one of the most complex. The complexity of the task is one of the explanations of why it is so difficult to use writing in order to express ideas, feelings or arguments (Graham & Harris, 1996). Creativity and productivity, both essential elements of writing, are noted to be lacking at school and in life. Although students and teachers recognize the importance of writing as an activity for improving thought, human intelligence and scholastic achievement, they also recognize that their skills as writers are quite poor (Shell, Murphy & Bruning, 1989).

According to Mata (1997), the field of writing can be structured along three themes: 1) the evaluation of compositions (the oldest and perhaps most developed); 2) the analysis of psychological processes involved in composition; and 3) the evaluation of didactic strategies for teaching composition. These themes in some way correspond to three approaches to teaching: writing as a product; writing as a process; and writing immersed in the context. Each of these approaches, from its particular standpoint, offers data and responses to the what, how and when of teaching and learning the writing process.

Writing as a process has been explained by contemporary cognitive models (Flower & Hayes, 1980; Scardamalia & Bereiter, 1992). These models describe the mental operations that take place when writing. In doing so, they study three subprocesses (planning, translating and reviewing), the differences between expert and novel writers and strategies for teaching and learning. The task of writing a coherent, adequate text involves several recurring stages during which the writer must juggle a set of specific procedures: 1) Planning (writing purpose, expected audience, content); 2) Composing or translating (characteristics of the type of text, adequate lexicon, correct morphosyntaxis, cohesion, spelling, punctuation); and 3) Reviewing (going over what has been written, rereading and evaluating it). This model from Linda Flower and John Hayes (1980) is quite useful for describing how thought and language are related in writing. The model defines writing as a problem-solving activity where three elements intervene and interact: the context of the task, long-term memory and working memory. In order to carry out the operations that make up writing, the person needs to use both working memory and long term memory. Working memory must be used as Hayes (1996) describes, where the writer needs to store the output from the stages of drafting,

expression and revision, as well as to make use of this information to direct the mental operations involved in each of these stages. This means that, while writing, the writer must make constant use of information from working memory. Similarly, as the writer executes the different stages of the writing process, he or she needs to use information stored in long-term memory: to evoke information for determining what content to include in the text when drafting, to make use of codes, different sources, resources and linguistic competencies during expression, or to direct the diverse actions that take place when revising.

Becoming aware of the phases involved in writing is what comprises the elements that uphold the modern theory of effective writing (metawriting). A person uses his or her awareness for writing when the writing process is carried: first, by remaining aware of the stages involved, and second, by staying alert to monitoring, evaluation and regulation of his or her writing performance.

Stages of writing

Understanding the stages of writing is fundamental to the development of meta-writing. The stages are: planning, translating (or expression) and revising.

1) *Planning* is a pre-writing activity, during which individuals conceive or generate the ideas that will make up the text that they wish to write, and then represent these in their mind. It is in this stage that the writer establishes the directionality of what he or she wishes to write.

2) *Translating* is the stage when the person writing makes use of his or her skills in order to produce the written version of the text; in other words, the text represented in their mind is translated into words, and then the words into their graphic expression (Galbraith & Torrance, 1999; Graham & Harris 2000). As a general position on what happens in this stage, we must clearly understand that a text *per se* has no meaning whatsoever. A text, to begin with, is a set of cues (codes), of varying nature, that guide the reader into constructing his or her own meaning.

3) In the final stage of *revising*, the writer focuses on rereading what has been written, keeping a critical eye, in order to adjust the parts that seem less than satisfactory. Revising allows the author of the text to use his or her own criteria in order to ensure several things:

first, that there is congruency between what is expressed in the text and the directionality that was intended; second, that there is cohesiveness among the different parts of the text, and between what is expressed and how the author thinks and feels; third, that the information presented is sufficient in quantity, there are no excesses or lacks, that what has been written is adequate and comprehensible to most of the intended audience, and the composition is sufficiently fluent, harmonious and elegant. Finally, the writer must ensure that the text is acceptable in terms of the language used; in other words, whether the language elements are used appropriately.

Metacognitive strategies in writing

But in order to elaborate a coherent, comprehensible written text, it is not enough to understand the stages to be followed when writing. It is also necessary to be aware that, the more we understand what we know, and the why behind our successes and shortcomings, the greater will be our chances for learning, consolidating successes, overcoming deficiencies and reducing our dependence on luck. This is what brings about meta-writing: being aware of what is being written and how it is being written, as well as overcoming faults by using metacognitive strategies. Based on the arguments proposed by Graves (1996) and Heller and Thorogood (1995), a structure has been established to classify metacognitive strategies for producing texts, or to consciously ascertain the goodnesses and limitations of thinking processes that give rise to a written text. These *suggested metacognitive strategies* are:

1) *Before writing*:
- Explain the objective one has in mind when writing.
- Construct ideas about the content of the text and its relationship to prior knowledge.
- Think about: who am I writing for?, what type of text do I want to write? why am I writing?

2) *During writing*:
- Identify difficulties that make it hard to express one's ideas.
- Use words, sentences and paragraphs to communicate.
- Ensure text coherence and cohesiveness.
- Select main and supporting ideas.

Electronic Journal of Research in Educational Psychology, *13*(3), 631-656. ISSN: 1696-2095. 2015, no. 37
http://dx.doi.org/10.14204/ejrep.37.15009

3) *After writing*:

- Analyze the quality of the content: scope, richness, depth and accuracy of the ideas.
- Review the organization and presentation of the content: text structure and the units it contains.
- Confirm or reject the writing style: suitability to the audience and purpose.
- Correct aspects of morphology, syntax, lexicon and spelling.

Understanding and constructing a written composition are tasks that demand strategic thinking on the part of the student. In other words, the writer must have the ability to master techniques and also know how to modify them to fit changing situations (Solé & Teberosky, 2010). Strategic thinking includes executive functions like the ability to inhibit automatic responses, problem-solving, task planning, working memory, strategy development, strategy implementation and being able to switch between strategies as the occasion requires (cognitive flexibility).

Executive functions may be defined as the set of cognitive skills that enable one to anticipate and set goals, to design plans and programs, to initiate activities and mental operations, to self-regulate and monitor tasks, be flexible in cognitive work and to organize it in time and space. All of this is prerequisite to behavior that is effective, creative and socially acceptable (Tirapu-Ustárroz & Muñoz-Céspedes, 2005). *Metacognitive skills* mainly refer to executive control processes. These procedures enable people to effectively link together the steps needed to successfully complete a task. They include skills for regulating cognitive processes, including the important skills of planning, monitoring and evaluation (Flavell & Wellman, 1977).

Finally, *metacognitive executive functions* are the set of cognitive and neurobiological processes that underlie the mental ability to internally prepare executive control. This control is at the base of self-reflection and self-awareness processes, needed for working out the final response that will effectively solve the problem. Metacognition has been approached from three types of knowledge: declarative knowledge (what knowledge or steps are required to carry out a task); procedural knowledge (how these steps can be carried out); conditional knowledge (the when and why of using one strategy and not another in order to execute a given task).

Metacognitive functioning in writing

According to Karmiloff-Smith (1995), a constant representational redescription of one's knowledge throughout the writing process suggests a model of evolutionary change. This model is based on the idea that the writer's knowledge becomes more and more explicit and accessible: the child starts with representations of a procedural nature until he or she is able to master them so as to explain and transform them into declarative representations. This postulate is very important in the area of metacognition, that is, the repetitive recoding of one's knowledge in long-term memory makes possible higher levels of metalinguistic awareness.

Based on the theory from Karmiloff-Smith (1995), researchers Flórez, Mondragón, Pérez and Torrado (2003) suggest three levels of metacognitive functioning in writing. They break down the characteristics of the three metacognitive processes (planning, self-regulation or monitoring, and evaluation) according to each level, and also take into account what knowledge the writer has about his/her own cognitive resources when approaching the written production task. Self-regulation or monitoring is identified with the drafting process, and evaluation with the revision process.

First-level or *Level I* metacognitive functioning in writing is procedural, where the writer's explicit understanding of his/her writing is still rough, with many overlooked details in the information being used in descriptions. Knowledge is handled implicitly (Karmiloff-Smith, 1995). At this level, planning is characterized by setting at least one specific objective and by choosing known strategies for successfully carrying out the task. The self-regulation process keeps the writer aware of the previously established objective, whereby he or she can correct the content, if needed, for that objective to be better understood. It is in this process that the writing itself takes place. Finally, the writer revises (evaluates) the content of what is written, thinking about the possible readers.

Second-level or *Level II* metacognitive functioning in writing refers to being able to bring to one's awareness the knowledge acquired, but this reflection is not always conscious, nor is it verbalized. This is an intermediate level. At this level, the planning process includes an implicit questioning about the characteristics that the text should have, and what one attempts to achieve with the composition. Next, the knowledge one needs is evoked, and some strategy is employed during writing in order to successfully complete the writing task; if

needed, the writer makes corrections to form, for better comprehension of the written text. Finally, the general characteristics of the text are identified and evaluated as to whether they meet the objective that was set earlier.

Metacognitive functioning in writing at the *third level*, or *Level III*, is the most complex. Knowledge acquired about writing is in long-term memory, so it can be brought into one's awareness and reflected on, analyzed and verbally expressed. For planning, the writer asks himself/herself about the characteristics of the text as well as his/her personal knowledge of the subject. The process of self-regulation makes the writer implement one general strategy and at least two more specific strategies. While writing, he or she uses prior knowledge, already evoked, and simultaneously identifies and corrects the written content in order to make the text more understandable. Once the task is finished, the written work is revised or evaluated, and changes are made to at least one aspect of the content in order to facilitate the reader's comprehension.

Objective

The aim of the present study is to examine the characteristics of EVAPROMES (Ulate-Espinoza, Jiménez-Rodríguez, Alvarado-Izquierdo & Puente-Ferreras, 2015), a test designed to assess writers' perceptions and writing skills as well as to identify schoolchildren with low and high writing awareness, among students in Spain and Costa Rica. The test evaluates processes that are at work in a creative writing task, giving less priority to mechanical writing, and focusing on meta-writing as the engine for change and improvement in the quality of written production. The diagnostic elements resulting from the test make it possible to define lines of action for the development of writing instruction and improvement programs.

Method

Participants

Participating in the study were a total of 1442 students, 915 Costa Ricans and 527 Spaniards, between the ages of 10 and 15. The sample from Costa Rica was a representative sample of sixth-graders, since a systematic random sampling procedure was followed, draw-

ing from 31 of the 155 schools that make up the *DREO* (Western region educational administration).

In Spain, the participating schools included public, subsidized and private schools, thereby reflecting the diversity and present state of schooling in the Madrid Region (56% public, 24% subsidized and 20% private). The sample was 527 students (51.9% boys, 48.1% girls), 60% of whom were sixth-graders, and the remainder from compulsory secondary education (7th to 10th graders). Spanish sample selection was made so as to match the ages represented in the Costa Rican sample, with a range from 10-15 years of age, and a mean age of 11.71 (SD= 1.17). Note that the age of sixth graders in Costa Rica, depending on the region, may correspond to the age of sixth or seventh graders in Spain.

Instruments

Composition or quality of the written production. A test was designed, consisting of free composition of a story or fantasy tale. A system of judges was used to assess the quality of the writing; 13 writing quality criteria were rated on a scale of 0 to 2. The procedure was tested and refined in successive applications, until there was practically unanimous agreement among the judges when rating the test compositions.

ESCOLA (scale of reading awareness; Puente-Ferreras, Jiménez-Rodríguez & Alvarado-Izquierdo, 2009). In order to assess the relationship with other metacognitive variables, and obtain convergent validity evidence, we used ESCOLA, a test designed to assess schoolchildren's level of metacognition in reading (see Jiménez-Rodríguez, Puente-Ferreras, Alvarado- Izquierdo & Arrebillaga-Durante, 2009).

Motivational climate in the classroom. In order to measure motivational climate, we used the CMC-Q questionnaire (classroom motivational climate), developed by Alonso-Tapia and Fernández (2008). The authors' research on classroom goal structure has demonstrated the need to assess motivational climate in order to facilitate evaluation of educational activities and promote their improvement. The psychometric characteristics of the CMC-Q are satisfactory, with high internal consistency values obtained in different samples, between .92 and .93.

Academic performance. An average was taken of the schoolchildren's grades for the year, in the following subjects: Social Sciences, Mathematics, English (EFL) and Spanish.

Procedure

In order to examine the theoretical bases of EVAPROMES, and its practical, conceptual and psychometric strengths and limitations, several tests were carried out in Costa Rica, specifically in the Western (DREO) and Alajuela (DREA) educational districts, and in Spain, in the Madrid region. We summarize the process carried out to validate the EVAPROMES scale, organized in two stages:

First stage (pilot version). The first version of EVAPROMES had a self-report scale design, containing 50 items; these were selected for their relevance and representativeness in relation to the three metacognitive processes previously described, and from the practical experience of teachers and researchers in interviews and focus groups. In addition, the theoretical and grammatical revision of this scale included assistance from expert researchers at the Complutense University of Madrid (UCM), Autonomous University of Madrid (UAM) and University of Barcelona (UB), with whom each scale item was discussed, in order to improve the scale.

The response design is a five-point Likert-type scale. This initial, pilot version was applied to an intentional sample, made up of 210 Costa Rican and 260 Spanish students in the fifth and sixth grades. Gender distribution was 63% boys and 37% girls, with ages between 10 and 15 years. Exploratory factor analyses pointed to a unidimensional structure with high reliability (Cronbach alpha of 0.92), and not allowing proper differentiation between the three reading processes. Furthermore, there were indications of biased responses due to social desirability, which was attributed to the use of a Likert type response scale. This type of scale makes it easy to associate high scores (on direct items) and low scores (on inverse items) as the most desirable choices.

Second stage (final version). The 28 items that had shown the best functioning in Stage 1 were selected, but now the response to each statement had to be made on a multiple choice basis, where the options (a, b, c) represented different strategies revealing different levels of metacognitive awareness. The final version of the instrument, along with the other instruments mentioned, were administered to broad, representative samples in both Costa Ri-

ca and Spain. The new instrument presented better differentiation between the reading processes, freedom from the social desirability bias, adequate validity evidence, and acceptable reliability, as shown in the Results sections.

Data analyses

Validity and reliability analyses were performed on scores obtained from both countries. As for validity, techniques from confirmatory factor analysis (CFA) and multi-group CFA were used to establish measurement invariance between countries, and other important variables such as gender and age. Evidence of convergent validity was also collected through measures of related constructs, such as ESCOLA and the motivation scale, CMC-Q. Finally, reliability through internal consistency was estimated using standard Cronbach alpha procedures, McDonald's omega procedures and the GLB (greatest lower bound of reliability), implemented with the Factor 9.3 program (Lorenzo-Seva & Ferrando, 2006).

Results

Structural analysis and construct validity

The theoretical foundations of EVAPROMES identify it as a complex measure, its design involving the incorporation of items that correspond to the metacognitive processes of Planning, Monitoring and Evaluation, applied to the variables Text, Task and Person. Despite the inherent complexity of the instrument, the first-stage analyses showed high reliability, which were confirmed in the second-stage analyses, as can be seen in the Results section. This reliability is in line with our objective to measure one global construct which we will call *metacognitive strategies applied to writing*, and three underlying, interrelated processes (planning, monitoring and evaluation).

EVAPROMES is composed of 28 items that are answered on an ordinal scale. The items were drafted following a design of 3 processes x 3 variables, resulting in nine groupings of items: (1) PLATEX items that measure planning applied to the variable text, (2) PLATAS items that measure planning applied to the variable task, (3) PLAPER items that measure planning applied to the variable person, (4) MONTEX items that measure monitoring applied to the variable text, (5) MONTAS items that measure monitoring applied to the variable task, (6) MONPER items that measure monitoring applied to the variable person, (7) EVATEX

items that measure evaluation applied to the variable text, (8) EVATAS items that measure evaluation applied to the variable task and (9) EVAPER items that measure evaluation applied to the variable person.

Through CFA, and applying robust maximum likelihood, goodness of fit was assessed for both the unidimensional and the three-factor model (Planning, Monitoring and Evaluation).

Table 1. **Goodness of fit of the models proposed for EVAPROMES**

Model	χ^2 (df)	Probab.	RMSEA	SRMR	NNFI	CFI
One factor	71.73 (27)	< .01	.04	.04	.94	.95
Three correlated factors	57.48 (24)	< .01	.04	.03	.95	.96
Three correlated factors with correlated errors	28.29 (21)	.13	.02	.02	.99	.99

Note: RMSEA = Root Mean Square Error of Approximation, SRMR= Standardized Root Mean Squared Residual; NNFI= Non-Normed Fit Index; CFI = Comparative Fit Index.

Table 1 shows that, while the fit to one factor (unidimensional model) is acceptable (RMSEA and SRMR less than .08 and NNFI and CFI at least .95), the three-dimensional structure represents a statistically significant improvement in terms of χ^2 =14.25, p< .01. In addition, a substantial improvement is gained in terms of χ^2 fit when introducing correlations between errors of the Text and Task variables (see Figure 1).

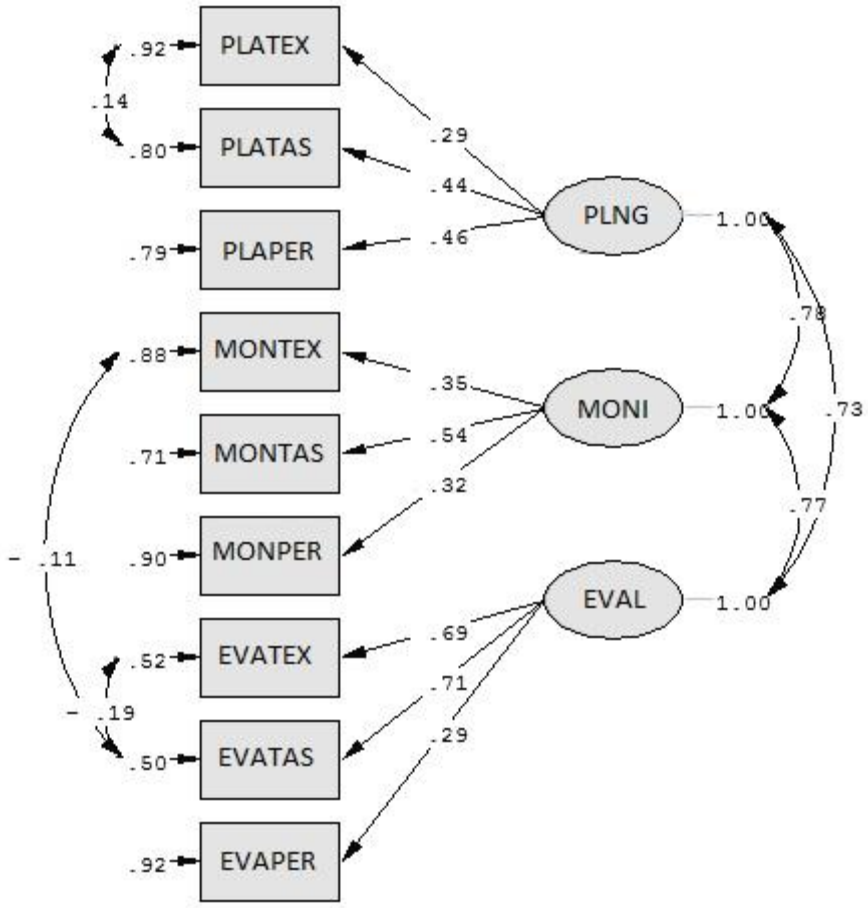

Figure 1. **Three-factor model of EVAPROMES**

Figure 1 illustrates the model that shows best fit, both statistically and substantively. On one hand, a mean between-factor correlation of .75 is adequate for considering that there is an important common variance, the measure of metacognition applied to reading, while at the same time we can discern between processes that form part of this measurement. The existence of overlaps (correlation between errors) between the contextual variables of Task and Text is explained in that both of them share common elements that are easily differentiated from the variable Person, or aspects related to it.

The model represented in Figure 1 was obtained from data from the Costa Rican sample, the broadest sample, although its goodness of fit is also observed in the Spanish sample $^{2}(3)= 13.70$ $p= .88$, RMSEA< .01. For a more systematic evaluation of invariance, we present below the results from multi-group CFAs where we checked for the test's metric equivalence in the different groups and populations studied.

Studies of metric equivalence and measurement invariance

In the published standards for elaborating psychological and educational tests, from the three large American associations (APA, AERA, NCME, 2014), a key aspect for test validation is being able to demonstrate that they meet the requirement of *fairness.* It is necessary to avoid biases in interpreting scores when diverse groups are compared; comparisons are unacceptable unless it is first shown that they are measuring the same construct, in other words, that there is metric equivalence or measurement invariance. There are different levels of invariance: the most basic is configuration invariance, verified when the same model fits the data from the different groups; a more demanding level of invariance is metric invariance, when the weights of items and their reliabilities are similar between groups; if the means also concur, we speak of measurement invariance; and the highest level of invariance is scalar invariance, when even the error variances are equal between groups.

Invariance Study between the Spanish and Costa Rican samples

Table 2 indicates goodness of fit indices of the multi-group CFAs, with estimate of robust maximum likelihood.

Table 2. **Multi-group CFAs with Spanish and Costa Rican samples**

INVARIANCE	χ^2 (df)	*p*	RMSEA	SRMR	NNFI	CFI
Configuration	53.36 (45)	.18	.02	.02	.99	.99
Metric	73.80 (51)	.02	.03	.04	.98	.99
Metric and covariance	84.10 (54)	.01	.03	.05	.98	.98
Scalar	266.35(57)	<. 01	.07	.04	.85	.88

The metric invariance model with equal covariances shows good fit, as can be seen in Table 2, even though its goodness of fit is slightly less than the configuration model and statistically better than the scalar model. Consequently, it may be stated that measurement of the construct *metacognition in writing* is invariant in both samples, analyzed at the configuration and metric levels, making it legitimate to do group comparisons with respect to the level of awareness in metacognition.

In both the Spanish and Costa Rican samples, large differences were observed between the different schools in the average levels of writing awareness, and these differences were related to how concerned the schools were with teaching metacognitive strategies. We

observed, in turn, that the schools most highly rated in Costa Rica and in Spain for their academic outcomes (CID tests given in sixth grade) were those that obtained the highest mean scores on EVAPROMES.

Invariance Study between genders

Table 3 indicates the goodness of fit indices for the different invariance models analyzed.

Table 3. **Multi-group CFA: boys vs. girls.**

INVARIANCE	χ^2 (df)	*p*	RMSEA	SRMR	NNFI	CFI
Spanish sample						
Configuration	50.54 (45)	.26	.02	.03	.99	.99
Metric	60.39 (51)	.17	.03	.05	.98	.99
Metric and covariance	63.13 (54)	.19	.03	.05	.98	.99
Scalar	113.54 (57)	<.01	.06	.04	.90	.92
Costa Rican sample						
Configuration	49.83 (45)	.29	.02	.03	.99	.99
Metric	59.21 (51)	.20	.02	.04	.98	.99
Metric and covariance	62.92 (54)	.19	.02	.04	.98	.99
Scalar	84.54 (57)	.01	.03	.03	.95	.96

In Table 3 we can observe good fit for both the configuration model and the metric invariance model, and even when covariances are fixed between the three processes. Statistically significant differences are only observed when comparing the previous models with the scale invariance model (this model no longer shows chi-square goodness of fit with $p< .05$ nor with values of NNFI< .95 or CFI< .95 in the Spanish sample). An absence of scale invariance was expected, since girls usually obtain higher scores on the verbal tests than boys, and this also occurs on EVAPROMES.

Invariance Study between Primary and Secondary Education. In the Spanish sample, in addition to the sixth graders, the test was also administered to seventh- and eighth-graders, allowing us to perform an invariance evaluation between primary and secondary education. (See Table 4.)

Table 4. **Multi-group CFA, Spanish sample: boys vs. girls.**

INVARIANCE	χ^2 (df)	*p*	RMSEA	SRMR	NNFI	CFI
Configuration	46.34 (45)	.42	.01	.04	1.00	1.00
Metric	44.58 (48)	.61	<.01	.04	1.00	1.00
Metric and covariance	49.35 (51)	.54	<.01	.06	1.00	1.00
Scalar	61.95(57)	.30	.02	.07	.99	.99

When comparing the sixth-graders with students from secondary education, scale invariance is obtained, thereby justifying application of the instrument across the age range of 10 to 15.

Sources of external evidence: predictive validity and relationship to other variables

In this section, external evidence is presented regarding the utility of EVAPROMES for predicting schoolchildren's academic performance and the quality of their written production. The relationship between EVAPROMES and other measures of metacognition will also be shown, as well as between EVAPROMES and variables related to motivational climate. Through an analysis of the correlation and regression patterns, we then show the relationship between EVAPROMES and the following measures and instruments: composition or quality of written production, ESCOLA, motivational climate and academic performance. For this study, a subsample of 349 sixth-graders were administered the tests indicated in Table 5.

Table 5. **Correlation patterns between metacognition tests and the criteria of Composition and Academic performance.**

	ESCOLA	Climate	Compo-sition	Social Sciences	Natural Sciences	Math	English as a foreign language	Spanish Language
EVAPROMES	.42***	.22***	.64***	.31***	.33***	.34***	.28***	.37***
ESCOLA	-	.25***	.41***	.31***	.36***	.29***	.31***	.34***
CLIMATE		-	.16**	.16**	.11	.14*	.10	.19**

***p< .001; **p< .01; *p < .05

We may observe in Table 5 that there is a medium size correlation between EVAPROMES and ESCOLA, and that the relationship with motivational climate, while significant ($p< .001$), is much weaker than expected. The correlation of .42 between EVAPROMES and ESCOLA indicates that there is a part of common variance between the two measures (approx. 18%). This is consistent with our intent to assess different expressions of metacognitive skills: with ESCOLA, reading awareness, and with EVAPROMES, writing awareness.

As for the predictive validity of EVAPROMES, we observe a correlation of .64 with the composition measurement (quality of written production), constituting evidence of its utility for predicting this criterion. In order to know which processes in EVAPROMES have the greatest predictive value for the composition measure, a regression analysis was performed (see Table 6).

Table 6. **Regression model for the three EVAPROMES measurements.**

	Non-standardized coefficients		Standardized coefficients	
	B	Std. Error		*t*
(Constant)	6.23	0.70		8.87*
PLANNING	0.20	0.04	.22	4.60*
MONIT.	0.36	0.05	.35	7.25*
EVAL.	0.28	0.04	.30	6.23*

*$p< .001$

The regression analysis shows that the three processes have predictive value at $p< .001$, accounting for 42% of the variance of written composition, corresponding to a multiple validity coefficient of R = .64. The fact that none of the processes is eliminated in the stepwise regression analysis due to possible overlap issues, or for *colinearity* (recall that the correlation between processes was .75), provides additional evidence that EVAPROMES measures three different processes and that each of them participates separately in predicting written production.

Prediction of academic performance is another possible application for metacognitive tests. As seen in Table 5, both EVAPROMES and ESCOLA show statistically significant correlations ($p< .001$) with performance in different subjects, most strongly in Spanish language: if both measures are taken as predictors of grades in Spanish, we obtain a predictive

validity coefficient of R= .43 (see Table 7). We find that the two tests, far from overlapping, are complementary, and together provide a better prediction of academic performance.

Table 7. **Regression model of metacognitive tests for predicting grades in Spanish Language.**

	Non-standardized Coefficients		Standardized coefficients	
	B	Std. Error		*t*
(Constant)	50.36	4.41		11.43*
EVAPROMES	0.49	0.10	.29	4.66*
ESCOLA	0.21	0.06	.23	3.73*

*p< .001

Reliability

EVAPROMES is an instrument containing items on an ordinal, three-point scale (0, 1, 2); in such cases it is recommended to check the asymmetry and kurtosis of the items' univariate distributions. If these are high, one should estimate the polychoric correlations matrix for later analyses, and estimate reliability using McDonald's omega coefficient or GLB, because the alpha coefficient usually underestimates reliability in these cases. Notwithstanding, GLB should only be used in medium- and large-sized samples (n > 500) in order to avoid overestimations of reliability.

In the Costa Rica sample, EVAPROMES reliability falls between .80 and .86 depending on the estimation procedure used (= .80, ω = .84 and *GLB* = .86). In the Spain sample, the reliability estimate according to this procedure ranges from = .80, = .86 to *GLB* = .91. Consequently, EVAPROMES reliability falls into similar ranges for both samples, with an average estimate of .85, which is an adequate value for carrying out between-subject comparisons.

Discussion and Conclusions

EVAPROMES is an instrument designed to assess meta-knowledge applied to writing, with instrument scores validated in two Spanish-speaking populations: Costa Rica and Spain. Analyses of scores obtained in the two samples suggest that EVAPROMES has excellent psychometric properties both in internal consistency reliability and in its validity. The structure of

the test, as in the theory underlying item definition, indicates three processes that are differentiated in EVAPROMES: planning, monitoring and evaluation, and that these processes unfold or are implemented as a function of three variables (Person, Task and Text). Therefore, while it is possible to obtain a global measure of performance, for the more precise assessment that is desirable for educational intervention, it is advisable to obtain a measurement for each process and variable. This way one can obtain a more precise profile of the student's metacognitive skills when writing.

One key aspect, often overlooked when creating a new instrument, is the need to obtain invariant measures; in other words, before applying an instrument intended to differentiate between groups, evidence is needed that the same construct is being assessed. In this sense, the present investigation has shown that EVAPROMES measures the same construct in two Spanish speaking populations that have quite different geographic and cultural characteristics. This is an indication that it might be further generalized, although this conclusion must be corroborated in later studies with samples from other Spanish speaking countries.

EVAPROMES has proven to be useful to predict the quality of schoolchildren's written production, but even more importantly, EVAPROMES and ESCOLA are shown to be related to academic performance and learning, thus confirming the educational importance of mastering metacognitive strategies (Coutinho, 2007; Dunning, Johnson, Ehrlinger & Kruger, 2003; Favieri, 2013, Kruger & Dunning, 1999). Self-regulation is a good predictor of academic performance in its different facets; use of self-regulation strategies like planning, monitoring, goal setting and perseverance are essential to academic achievement in schoolwork (Álvarez-Valdivia, 2009; Corno, 1986; Martínez-Vicente & de la Fuente, 2004, Pintrich & De Groot, 1990; Zimmerman & Pons, 1986, 1988), as well as to the student's own self-assessment throughout the learning process (Panadero & Alonso-Tapia, 2013). Correlational and linear regression studies allowed us to evaluate the relationship between the use of metacognitive strategies applied to writing, and other related constructs such as meta-knowledge in reading (ESCOLA) and motivational climate in the classroom (CMC-Q), quality of written texts and academic performance. We conclude that EVAPROMES has proven to be a highly capable instrument for predicting the quality of schoolchildren's written production, and along with ESCOLA, to act as a good predictor of academic performance in two basic instrumental areas (reading and writing). We recommend the use of both instruments, whenever

possible, in order to design more effective teaching methodologies that reduce school failure, and, when needed, to make a better diagnosis for carrying out more effective interventions.

Educators have taken an interest in writing primarily as a tool for cognitive development; for this reason, it is important to understand the metacognitive strategies that foster writing skills: a) Make students write frequently; few scholastic activities produce more cognitive development than making students write often and providing feedback on their writing (Calkins, 1994); b) Create an informal, supportive environment for writing; this promotes spontaneous, creative writing (Englert, Raphael, Anderson, Anthony & Stevens, 1991); c) Strengthen pre-writing strategies as a condition needed for organizing its execution (Bereiter & Scardamalia, 1987); d) Emphasize transformation of knowledge rather than narrating it (Beal, 1996); e) Encourage the student to develop productive revision strategies; dialogue with the peer group can improve revision strategies (Englert et al., 1991); f) Use computer technology when needed, the use of computers can particularly facilitate the writing of quality stories (Berninger, Vaughan, Abbott, Abbott, Rogan, Brooks, Reed & Graham, 1997).

Writing skill is developed slowly in most children and adolescents, and only reaches maturity in adulthood -- and even then, only for a few writers (Bartlett, 2003). There is great concern in schools today about the number of children who do not write or who write poorly. One study from the US described writing as "the Neglected R" (National Commission on Writing, 2003). Although many primary and secondary students come to master the basic elements of writing, only a small number, about 25%, write competently. To write competently involves good organization, with "convincing and elaborated responses to the tasks assigned, and the use of rich, evocative and compelling language" (pp. 16-17). In order to improve children's writing, the commission recommends that schools double the time dedicated to writing; and that writing form part of all school subjects.

Limitations and future research

One of the limitations we observe with EVAPROMES is the age range that it addresses. Instruments need to be created that can assess metacognitive processes in writing throughout secondary education and even in university students. We have observed that pre-university and university students in Spain do not apply metacognitive strategies to writing tasks, we thus consider it of interest to assess what processes and variables are put into play when they are writing. If we obtain a metacognitive diagnosis of these students, we will be

able to implement *ad hoc* programs to correct deficiencies and thus improve their academic performance.

Although this test has worked well in two Spanish speaking countries, it would be helpful to study how the instrument behaves in other countries where Spanish is spoken. Future research should include samples from other countries in order to analyze possible generalization of results to the Spanish speaking population, In this way the test might be established internationally and become a useful tool for diagnosis and intervention in the Spanish language.

References

Alonso-Tapia, J. & Fernández, B. (2008). Development and initial validation of the Classroom Motivational Climate Questionnaire (CMC-Q). *Psicothema*, *20*(4), 883-889. Retrieved from http://www.redalyc.org/articulo.oa?id=72720456

Alvarez-Valdivia, I. M. (2009). Assessment for contributing to self-regulated learning. *Electronic Journal of Research in Educational Psychology*, *7*(3), 1007-1030.

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (AERA, APA, NCME, 2014). *Standards for educational and psychological testing.* Washington, DC: American Psychological Association.

Bartlett, T. (2003). Why Johnny can't write, even though he went to Princeton. *Chronicle of Higher Education*, *49*, A39. Retrieved from http://chronicle.com/free/v49/i17/17a03901.htm

Beal, C. R. (1996). The role of comprehension monitoring in children's revision. Educational Psychology Review, *8*, 219–238

Bereiter, C. & Scardamalia, M. (1987). *The psychology of written composition*. Mahwah, New Jersey: Erlbaum.

Berninger, V. W., Vaughan, K. B., Abbott, R. D., Abbott, S. P. Rogan, L. W. Brooks, A., Reed, E. & Graham, S. (1997). Treatment of handwriting problems in beginning writers: Transfer handwriting to composition. *Journal of Educational Psychology, 89*, 652-666, doi: 10.1037/0022-0663.89.4.652

Calkins, L. M. (1994). *The act of teaching writing*. Portsmouth, NH: Heinemann.

Corno, L. (1986). The metacognitive control components of self-regulated learning. *Contemporary Educational Psychology, 11*, 333-346, doi: 10.1016/0361-476X(86)90029-9

Coutinho, S. A. (2007). The relationship between goals, metacognition, and academic success. *Educate~, 7* (1), 39-47.

Dunning, D., Johnson, K., Ehrlinger, J. & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, *12*(3), 83-87, doi: 10.1111/1467-8721.01235

Englert, R. S., Raphael, T., Anderson, L., Anthony, H. & Stevens, D. (1991). Making strategies and self-talk visible: Writing instruction in regular and special education classrooms. *American Educational Research Journal, 28*, 337-372.

Favieri, A. G. (2013). General Metacognitive Strategies Inventory (GMSI) and the Metacognitive Integrals Strategies Inventory (MISI). *Electronic Journal of Research in Educational Psychology*, *11*(3), 831-850. doi: http://dx.doi.org/10.14204/ejrep.31.13067

Flavell, J. H. & Wellman, H. M. (1977). Metamemory. In R. V. Kail Jr. & J. W Hagen (Ed.). *Perspectives on the development of memory and cognition* (pp. 3-33). Hillsdale, N. J.: LEA.

Flórez, R., Mondragón, S.P., Pérez, C. & Torrado, M.C. (2003). Explorando la metacognición: evidencia en actividades de lectura y escritura en niños y niñas de 5 a 10 años de edad. [Exploring metacognition: evidence in reading and writing activities in children between the ages of 5 and 10.] *Revista Colombiana de Psicología*, *12*, 85–98.

Flower, L. & Hayes, J. (1980). The dynamics of composing: making plans and jiggling constraints. In L. W. Gregg & E.R. Steinberg (Ed.). *Cognitive processes in writing.* (pp. 31-50). Hillsdale: LEA.

Galbraith, D. & Torrance, M. (1999). Conceptual processes in writing: from problem solving to text production. In M. Torrance & D. Galbraith (Ed.). *Knowing what to write: conceptual processes in text production.* Amsterdam: Amsterdam University Press.

Graham, S. & Harris, K.R. (1996). Self-regulation and strategy instruction for students who find writing and learning challenging. In C. M. Levy & S. Ransdell (Ed.). *The science of writing: theories, methods, individual differences and applications*. (pp. 347-360). Mahwah, N.J.: Earlbaum.

Graham, S. & Harris, K.R. (2000). The role of self-regulation and transcription skills in writing and writing development. *Educational Psychologist, 1*(35), 3-12, doi: 10.1207/S15326985EP3501_2.

Graves, D. (1996). *Didáctica de la escritura*. [The didactics of writing.] Madrid: Morata.

Hayes, J. (1996). A new framework for understanding cognition and affect in writing. In M.C. Levy & S. Ransdell (Ed.). *The science of writing*: *theories, methods, individual differences and applications*. (pp. 1-27). Mahwah, N.J.: Erlbaum.

Heller, M. & Thorogood, L. (1995). *Hacia un proceso de lecto-escritura reflexivo y creativo*. [Towards a reflexive, creative reading-writing process.] Caracas: Educativa.

Jiménez-Rodríguez, V., Puente-Ferreras, A., Alvarado-Izquierdo, J. M. & Arrebillaga-Durante, L. A. (2009). Measuring metacognitive strategies using the reading awareness scale: ESCOLA. *Electronic Journal of Research in Educational Psychology, 7*(2), 779-804.

Karmiloff-Smith, A. (1995). Restricciones de la conciencia metalingüística. [Restrictions of metalinguistic awareness.] *Infancia y Aprendizaje, 72,* 33-50.

Kruger, J. & Dunning, D. (1999) Unskilled and unaware of it: How differences in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*(6), 1121-1134, doi:10.1037/0022-3514.77.6.1121

Lorenzo-Seva, U. & Ferrando, P.J. (2006). FACTOR: a computer program to fit the exploratory factor analysis model. *Behavior Research Methods, 38*(1), 88-91, doi: 10.3758/BF03192753

Martínez-Vicente, J.M. & de la Fuente, J. (2004). Self-regulation of learning through the PRO & REGULA program. Electronic Journal of Research in Educational Psychology *2*(1), 145-156.

Mata, F. (1997). *Dificultades en el aprendizaje de la expresión escrita*. [Difficulties in learning written expression.] Málaga: Aljibe.

National Commission on Writing in America's Schools and Colleges (2003). *The neglected* New York: The College Board.

Panadero, E. & Alonso-Tapia, J. (2013). Self-assessment: Theoretical and Practical Connotations. When it Happens, How it is Acquired and what to do to Develop it in our Students. *Electronic Journal of Research in Educational Psychology, 11*(2), 551-576 doi: http://dx.doi.org/10.14204/ejrep.30.12200

Pintrich, P. R. & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*(1), 33-40.

Puente-Ferreras, A., Jiménez-Rodríguez, V. & Alvarado-Izquierdo, J. M. (2009). *ESCOLA. Escala de conciencia lectora.* [Reading awareness scale.] Madrid: EOS.

Scardamalia, M. & Bereiter, C. (1992). Dos modelos explicativos de los procesos de composición escrita. [Two explanatory models of written composition processes.] *Infancia y Aprendizaje*, *58*, 43-64.

Shell, D. F., Murphy, C. C. & Bruning, R. H. (1989). Self-efficacy and outcome expectancy mechanisms in reading and writing achievement. *Journal of Educational Psychology*, *81*, 91-100, doi: 10.1037/0022-0663.81.1.91

Solé, I. & Teberosky, A. (2010). La enseñanza y el aprendizaje de la alfabetización. [Teaching and learning literacy.] In C. Coll, J. Palacios, & A. Marchesi (Ed.). *Desarrollo psicológico y educación*. (pp. 462- 485). Barcelona: Alianza.

Tirapu-Ustárroz., J. & Muñoz-Céspedes, J.M. (2005). Memoria y funciones ejecutivas. [Memory and executive functions.] *Revista de Neurología, 41*(8), 475-484.

Ulate-Espinoza, M.A., Jiménez-Rodríguez, V., Alvarado-Izquierdo, J.M. & Puente-Ferreras. A. (2015). *EVAPROMES. Evaluación de los procesos metacognitivos en escritura*. [EVAPROMES. Assessing metacognitive processes in writing.] Madrid: EOS.

Zimmerman, B. & Pons, M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal, 23,* 614-628, doi: 10.3102/00028312023004614

Zimmerman, B. & Pons, M. (1988). Construct validation of a strategy model of student self-regulated learning. *Journal of Educational Psychology, 80*, 284-290, doi:10.1037/0022-0663.80.3.284

[This page intentionally left blank]

 Electronic Journal of Research in Educational Psychology, *13*(3), 631-656. ISSN: 1696-2095. 2015, no. 37
http://dx.doi.org/10.14204/ejrep.37.15009