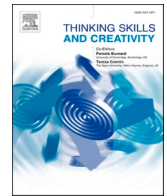




ELSEVIER

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

## Thinking Skills and Creativity

journal homepage: [www.elsevier.com/locate/tsc](http://www.elsevier.com/locate/tsc)

## Design and validation of CRISENSE, a novel critical competence assessment tool for Spanish adolescents and young adults

Elena del Pilar Jiménez Pérez<sup>a</sup>, Manuel León Urrutia<sup>b,\*</sup>, Pedro García Guirao<sup>a</sup>,  
María Victoria Cerezo Guzmán<sup>a</sup>

<sup>a</sup> Universidad de Málaga, Spain

<sup>b</sup> University of Southampton, UK

### ARTICLE INFO

#### Keywords:

Critical competency  
Competency test  
Validation tool  
Critical thinking  
Reading competence  
Confirmatory factor analysis

### ABSTRACT

The acquisition of critical competence is essential at all educational levels. The measurement of this competency presents a significant challenge, particularly considering the methodological complexities involved.

Developing tests to assess critical competence requires rigorous processes of elaboration and validation, without which the validity of the results can be affected. Most existing Spanish tests are translations from other languages –such as English. This often affects the readability of the questions, and in turn the validity of the tests. To address this issue, this study proposes and validates CRISENSE, a critical competency measurement tool created in Spanish that addresses three dimensions: comprehension, deduction, and critical positioning.

CRISENSE was designed based on the existing Watson-Glasser and Halpern critical thinking tests, and pre-tested by a panel of experts in critical competence and reading comprehension. The resulting tool was subjected to testing with a sample population of 575 students, comprising individuals from secondary, high school, and university levels of education. The psychometric properties of the test were evaluated through descriptive analysis, item homogeneity, confirmatory factor analysis, and a measurement invariance study between sexes and between educational levels. Additionally, the reliability of the scores was evaluated using Cronbach's Alpha coefficient. The results demonstrated positive evidence of validity and reliability of the test scores, exhibiting satisfactory adjustment indices and coefficients in each analysis. Consequently, the critical thinking test can be deemed as a valuable tool for assessing this construct in Spanish students, contributing to the expansion of current scientific evidence on this topic.

### 1. Introduction

The current age of information brings with it a deluge of misinformation. It is therefore imperative to establish a base for sharing and producing knowledge with a core critical thinking element (Nieto et al., 2009). Similarly, developing critical competence in all aspects of life is essential for healthy coexistence. For example, a truthful perception of what we read must be mediated by our own emotional management (Jiménez-Pérez, 2017) as well as the ability to make decisions as citizens based on knowledge and common sense.

\* Corresponding author.

E-mail address: [m.leon-urrutia@soton.ac.uk](mailto:m.leon-urrutia@soton.ac.uk) (M. León Urrutia).

<https://doi.org/10.1016/j.tsc.2024.101708>

Received 6 March 2023; Received in revised form 21 November 2024; Accepted 21 November 2024

Available online 26 November 2024

1871-1871/© 2024 The Author(s).

Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Published by Elsevier Ltd. This is an open access article under the CC BY license

The European Commission has recently placed special emphasis on critical thinking and creative competences in the Higher Education space (European Commission, 2021). It is worth noting that such competences already involved the need to cultivate a wide range of skill in university students, including the capacity for analysis and synthesis, the ability to manage information, problem solving, decision making, critical capacity, or the ability to generate new ideas. Moreover, the development of thinking skills in the educational field is a priority for many other public bodies (LOE, 2006; UNESCO, 1995, 1998, 2009). There is a consensus among researchers in the fields of philosophy, psychology, and education where critical thinking encompasses a range of skills, including analysis, logical reasoning, judgment, and decision-making (Gravet & Hainselin, 2021). All these competencies are fundamental to the pedagogical process of pupil education and knowledge acquisition (Blakemore, 2018), as well as essential for the exercise of global citizenry (Jiménez Pérez, 2023; Betancourth et al., 2017). Consequently, a critical perspective of reality can serve as the foundation for effective communication from an early age, such as the reading processes in primary education. A critical examination of the world can serve as the foundation for proficient expression, beginning in childhood and exemplified in the reading activities of elementary schooling (Gutiérrez-Fresneda, 2022).

Moreover, the mere act of teachers exposing students to content in the classroom does not guarantee that students will learn how to handle such knowledge appropriately (Redondo Gutiérrez et al., 2017). The objective is not to indoctrinate or dictate how reality should be understood. Rather, knowledge should become the foundation of freedom to make informed decisions based on reasoned judgment and optimal critical thinking; this enables students to regulate their processes (Rodríguez et al., 2012). Nevertheless, prior to attempting to educate students, it is essential to ascertain whether educators have undergone training to foster critical competence in their pupils (Izarra et al., 2020).

### 1.1. Critical thinking

The roots of critical thinking can be traced back to antiquity. In classical Greece, seminal philosophers such as Socrates, Plato and Aristotle developed their theories of *maieutics*, *dialectics* and *rhetoric*, respectively. In this vein, it draws on disciplines such as logic, which is grounded in common sense, reason and objectivity. To reach fully comprehension, it is also necessary to examine its etymology. Accordingly, the term “critique” has its etymological roots in ancient Greek, where it means one’s capacity for discernment. This etymological lineage is shared with other terms denoting similar faculties, such as “judge” and “judgment” (*kritēs*, “judge” and *krinō*, *krinein*, “to judge”).

According to Bachelard (1994), humans possess an inherent capacity for critical thinking, associated with their cognitive abilities. These include the ability to perceive, infer and process information (Puche, 2000). Such ability enables individuals to perform their learning function in an educational setting from a formal perspective (Spelke, 1991; Puche, 2000). For this reason, it is imperative to cultivate this capacity from an early age in the classroom, where opportunities are provided for students to self-regulate to enhance their efficiency and enable them to plan, monitor, and assess their teaching and learning independently (Al-Ahmadi, 2008; Tamayo et al., 2014). At present, studies in the field of education address the development of informal argumentation in text commentary as a requirement for teacher training (Caro Valverde et al., 2018; Vicente-Yagüe et al., 2019; Vicente-Yagüe Jara et al., 2019).

Although most aspects of critical thinking are commonly agreed among academics, there are certain variations in their conceptualization. While Mockus (1989) stated that encouraging rational discussion, the written tradition, and the restructuring of action in educational centers are fundamental elements of critical thinking, Bachelard (1994) argued that critical reasoning is predicated on the deployment of common sense, intuition, the projection of images and metaphors, parallelisms, and generalizations. For Kurland (2000), critical thinking entails adhering to the evidence, embracing the full spectrum of reality, placing greater reliance on reason than on intuition, maintaining accuracy even when it entails embracing the unpopular, and pursuing truth over the pursuit of correctness (being aware of one’s own limitations and biases). Similarly, Facione (2007) cites the Delphi report (1990), which defines critical thinking as an intellectual process that, in a structured and self-regulatory manner, strives to reach a reasonable judgment or conclusion. This is characterized by being the product of an effort of interpretation, analysis, evaluation, and inference from evidence. It can be explained or justified by the evidential, conceptual, contextual, and criteria considerations on which it is based. Paul and Elder (2003) define critical thinking as the process by which a person improves their thinking quality, whereas Seligman (2005) characterizes it as thinking calmly while contemplating all possibilities, while González Zamora (2006) views it as the ability to make good judgments in search of the truth.

The term “critical thinking” is often associated with concepts such as to reason, intelligence and intellect, which are believed to be inherent to human cognitive abilities. This association is so pervasive that some scholars have equated critical thinking with scientific thinking (Tamayo, 2009; Caravita & Hallden, 1994). For decades, a multitude of scientific and pedagogical frameworks have been established for critical thought. The Delphi study (1990) established critical thinking as a cognitive procedure geared towards reaching well-founded conclusions through self-directed means. In his 1994 work, Bachelard proposes that critical thinking should be grounded in several different facets, including common sense, intuition, imagined metaphors and parallelisms, among related facets. Kurland (2000) elaborates that this entails precisely surveying proof under the broad lens of reason, albeit with less consideration of intuition, and with an openness to entertaining contrary or unconventional views. As posited by Baghana et al. (2020), “it represents a student’s faculty to reflect more logically and lucidly”; this notion, however, is overly simplistic. Furthermore, passion has even been proposed as a key element within conceptualizations of critical thought (Sharma et al., 2022), potentially as a proxy for motivation. However, critical thinking remains, by definition, objective at its core. Consequently, it is becoming a 21st-century skill, designed to enable students to think for themselves, which is crucial for complex, globalized, and increasingly digitized economies and societies (Vincent-Lancrin et al., 2019). Nevertheless, even though educators and policymakers view critical thinking as a fundamental educational goal, there is still a lack of clarity surrounding the ways in which these abilities can be fostered within the school

environment (Palma Luengo et al., 2021).

It should be noted that critical thinking and critical competence are not necessarily synonymous. Critical thinking is the capacity of an individual to comprehend reality in a detached and impartial manner, subjecting it to objective analysis and evaluation. Critical competence, on the other hand, denotes the application of critical thinking in a meaningful and productive manner within one's own context, ultimately leading to informed decision-making (Jiménez Pérez, 2023). In other words, critical thinking represents a prerequisite for critical competence, which encompasses decision-making and action. Consequently, critical competence represents the foundation upon which contemporary educational practices should be built. This foundation encompasses four key areas: learning to know, learning to do, learning to live together, and learning to be. The concept of critical thinking serves as the foundation for Howard Gardner's *Five Minds for the Future* (2005) and Ken Robinson's essence of *The Element* (Robinson and Aronica, 2009).

To a certain extent, critical thinking can be identified within the framework of the international assessments of reading literacy, as exemplified by PIRLS (Mullis & Martin, 2019) and PISA (OECD, 2018). The final processes or aspects of those assessments are designed to assess quality and credibility, reflect on content and form, construct meaning, and develop personal potential (Marcano Medina, 2023; García Medina et al., 2020). It is evident that a negligible proportion of students in participating countries can attain such scores, particularly in light of the scores they have obtained. However, as has already occurred with the concepts of reading comprehension and reading competence, the term "critical thinking" has been employed, which is currently beginning to be supplanted by "critical competence". Nevertheless, the limitations and relationships between these two concepts remain undefined.

### 1.2. Critical thinking tests

While numerous questionnaires have been developed to assess critical thinking, there is currently a lack of instruments designed to evaluate critical competence. Furthermore, the tests of critical thinking in Spanish are primarily translations or adaptations of tools originally created in English. The challenge of identifying an optimal translation can negatively impact on the tool's outcomes (Fernández et al., 2021). This is because the ability to comprehend the author's intent and meaning, which is contingent upon reading proficiency, can influence the accuracy of the results (Jiménez-Pérez et al., 2021; Gutiérrez-Fresneda et al., 2017). Moreover, the specific dimensions that each author establishes vary across tests. Some consider the semantic-verbal-linguistic condition (HCTAES or CCTT) to be a form of critical competence, rather than a component of reading competence, that precedes the evaluation of criticism. Considering their pervasive utilization, three tests merit particular attention:

Firstly, the HCTAES test (Halpern, 2012) is a translation from English that evaluates five skills: Hypothesis Testing, Verbal Reasoning, Argument Analysis, Probability and Uncertainty, and Problem Solving and Decision Making. A total of 25 situations that emulate real-world scenarios are utilized, with double-format questions that are open and closed. The test is sufficiently challenging to last up to two hours. In contrast, the *PRENCRISAL* and *PENTRASAL*, which are presented as enhancements to the HCTAES, are tests comprising 35 open-ended questions on daily routines across diverse domains, with a single response. The time required for completion is one hour, and the number of questions is greater, which may prove tedious and repetitive for high school students.

Secondly, the Cornell Critical Thinking Test (CCTT) is designed to assess critical thinking abilities in individuals between the ages of 9 to 18 years. It employs multiple-choice format at Level X, comprising a total of 66 items. However, this number is excessive for students in elementary school, as the test is intended for older age groups. Furthermore, it assesses a range of critical thinking skills, including induction, source credibility, observation, semantics, deduction, and hypothesis identification, integrated with reading comprehension items.

Thirdly, the California Critical Thinking Skills Test (CCTST) is designed to assess the evaluation, interpretation, analysis, explanation, and inference abilities of the test-taker. It establishes six scores: a total score for the cognitive skills of critical thinking and five sub-scores for each of those dimensions. However, it is only intended for use with adults and university students. The Watson-Glaser III test, arguably the most renowned in education and recruitment, encompasses five dimensions: inference, recognition of assumptions, deduction, interpretation, and evaluation of arguments. The test comprises 40 items selected at random from a battery of questions of the Likert scale type questions, with an approximate duration of thirty minutes.

Those tests have certain limitations. First, they are designed to assess cognitive abilities, rather than practical skills. Second, the perspective is unduly focused on linguistics. Although linguistics is a significant element, this suggests that reading competence is a prerequisite, which does not align with the reality of individuals who are illiterate but possess advanced critical thinking abilities. Third, such limitations may be correlated with age; for example, the reading experience of an 11-year-old adolescent may be markedly different from that of a 23-year-old, introducing another layer of limitation. To address this, *CRISENSE* is designed to be read or listened; its purpose is to be developed as an app. Also, this is the inaugural attempt to delineate the fundamental parameters that define the process of exercising sound judgment and critical thinking in a practical manner. The three-step process is as follows: Observe, Think, and Act. In a developed society where timesaving is a primary concern, the phrase "act without thinking" assumes a particular significance.

## 2. The dimensions

Halpern differentiates between the following cognitive processes: hypothesis testing, verbal reasoning, argument analysis, probability and uncertainty, decision making, and problem solving. Similarly, he reflects the critical thinking approach of *PRENCRISAL* and *PENTRASAL*. In contrast, Cornell (CCTT) places greater emphasis on induction, source credibility, observation, semantics, deduction, and hypothesis identification. (Watson and Glaser, 1991), in contrast, identify five dimensions: inference, assumption recognition, deduction, interpretation, and argument evaluation.

In this sense, in all these cases these dimensions are subjective criteria that have not been previously validated by a panel of experts. Rather, they have been established as valid a posteriori through statistical analysis. These dimensions align with an exclusively psychological and cognitive approach that does not consider a more complex and extensive process. However, this approach is more logical and habitual. For instance, critical thinking is a cognitive process that can be applied to our daily life to enhance our decision-making abilities and facilitate our day-to-day routines. Consequently, critical thinking encompasses the process of comprehension, assumption, and deduction as a cognitive dimension through which reality is observed. In a second dimension, the analysis, evaluation, and criticism of information allows for the extraction of information and the generation of a passive, theoretical reaction to this reality is generated. Ultimately, the capacity to utilize the insights gleaned from a critical analysis of reality to inform decision-making and act in alignment with one's beliefs represents the defining feature of competence. In this way, we find ourselves confronted with the notion of the coherence that Einstein alluded to in his statement: "If you want to know a person, do not listen to their words but observe their actions."

To further the domain of pedagogy, it is essential to establish clear and robust directives concerning critical thinking and proficiency, along with the interconnection between these two elements. Similarly, human maturation within one's habitat, namely cohabitation as a citizen capable of making informed decisions based on well-developed, sound, and logical critical thinking, is essential. Considering the conceptual model of Jiménez Pérez (2023) on critical competence and the delineated facets (Observe, Think and Act), recommendations have been submitted to the INEE (National Institute for Educational Assessment) of Spain for the European Learning Zone. This assessment mechanism has been formulated for utilization in academic and professional settings. Therefore, this is the inaugural tool that actualizes the transition from critical thinking to critical competence (Jiménez Pérez, 2023) in a concrete manner, as one of the pivotal competencies demanded by Europe in accordance with its recommendations on lifelong learning (2018), which also pertains to the EEA (European Education Area).

The aim of this research is to validate CRISENSE, a tool that, as a novel contribution to the field, has been developed in Spanish to evaluate critical competence. Critical competence can be defined as the ability to use critical thinking in a useful way in the environment (Jiménez Pérez, 2023). The tool has been designed to assess this competence in adolescents and adults, from high school to university, in a time-efficient manner (approximately thirty minutes). It has also been developed to assess the dimensions of critical competence that have been previously validated by a panel of experts. These three dimensions are: a) To understand/suppose/deduce; b) To analyze/evaluate/criticize and c) To decide/conclude/position oneself (Jiménez Pérez, 2023); dimension "c" is defined as acting, which transforms thinking into a competence, representing a significant innovation.

### 3. Method

#### 3.1. Design of the questionnaire

The dimensions of the CRISENSE test to be assessed were identified through focus groups with professionals in the field of education and psycho-pedagogy. Firstly, a series of questions were generated, each framed within the three dimensions collected: understand/suppose/deduce, analyze/evaluate/criticize and decide/conclude/position oneself. These dimensions were used as a starting point for the construction of the items, with the models of Halpern and Watson, and Glaser, the most used, from a critical perspective, serving as the foundation for the development of items. In the construction of the items, the recommended writing rules were considered, namely clarity, representativeness, understandability, relevance, specificity and simplicity (Muñiz & Fonseca-Pedrero, 2019).

Subsequently, the items were subjected to an eight-panel review by experts from various Spanish universities and research centers, including the universities of Sevilla, Cádiz, Murcia, Granada, and the *Spanish Association of Reading Comprehension* (AECL in Spanish). This panel devised the initial questionnaire, which was modified to align with the three identified dimensions and recommended refinements to the phrasing of the questions, which were incorporated. Accordingly, the final expert-revised test comprised six brief texts, each based on three questions. Each of the three questions pertaining to the texts addressed one of the three dimensions: understanding/supposing/deducing, analyzing/evaluating/criticizing, and deciding/concluding/positioning (Jiménez Pérez, 2023). To illustrate, question 1, "You believe he is either impolite or holds you in contempt," pertains to dimension 1, "understanding/supposing/deducing." Question 2, "You concur with your acquaintance that the young man is undoubtedly a *protégé*," pertains to dimension 2, "analyzing, evaluating, and criticizing," while question 3, "This is not an auspicious beginning to the day; I would prefer to concentrate on what is essential," addresses dimension 3, "deciding, concluding, and positioning." The probability of developing a friendship was minimal, as indicated by dimension 3, "deciding/concluding/positioning." The subsequent triplets of questions, each with its own scenario, followed the same structural pattern. The questions were presented with multiple-choice options, allowing respondents to select a single response. The options remained consistent across all questions, employing a five-level scale ranging from "true" to "false" in relation to the statement presented in each question. The available responses were "true," "very likely," "likely," "unlikely," and "false." Each question was accompanied by a definitive answer, presented as either "true" or "false." The mean values of the keys did not fall within the scale. Scores were assigned in inverse proportion to the distance between the key and the answer. In other words, if the key was "false," the answer "false" was assigned five points, the answer "unlikely" four points, and so on down to the minimum score of 1, which consisted of the maximum distance to the answer key.

An example of questions by dimensional in this test is based on reading or listening to a short text (about 150–200 words, short sentences, objective data) with a Likert scale question for each dimension (observe, think, act). A situation is proposed in which different elements provide information about a fact that must be considered and reflected upon before choosing an answer for the decision-making. For example, in a text in which a worker claims to have suffered harassment at work, there are three signals in the

text that indicate whether this may or may not be the case (observe question), there are several options to choose from according to the law (think question), and finally, based on your conclusions you either show solidarity or not (act question).

### 3.2. Procedure

The ethical standards set forth in the Declaration of Helsinki were adhered to, as were the general ethical guidelines of the AECL (*Asociación Española de Comprensión Lectora* or Spanish Association of Reading Comprehension) for research projects employing questionnaires. These guidelines encompass the assurance of information rights and the protection of personal data, the provision of confidentiality guarantees, and the guarantee of non-discrimination for any reason. Additionally, they entail the provision of free, voluntary, and reversible participation. Each participant was informed of the objective of the study and provided with written consent. In the case of minors, their legal guardians were also informed and agreed to participate in the study, signing the aforementioned consent.

The questionnaire was designed to be completed in an electronic format, with all responses to the questions required. This indicates that no records were omitted. Responses from participants who did not complete the questionnaire within the specified timeframe were excluded from the analysis.

The present study was designed with a longitudinal quasi-experimental approach and was conducted in public secondary schools, high schools, and universities across multiple regions of Spain. The total number of students included in the sample was 575. The sample dataset has been shared in the repository of the University of Southampton (Jiménez-Pérez et al., 2024)

The data collection process was designed in collaboration with the teaching staff at the aforementioned educational institutions, who were contacted via email through the AECL database. The questionnaire was completed by students at the participating educational centers during school hours via the Internet at the following link: <https://forms.gle/3AXEawn9HmpQrodKA>. The time allotted for completion of the test was half an hour. This was done to avoid critical hours such as the first and last periods, and to ensure that the tests were completed in a distraction-free environment. Tests were conducted on individual electronic devices, and in total silence to avoid mutual feedback.

### 3.3. Participants

The sample was composed of secondary school, high school, and university students, resulting in a total population of 575 participants. The participants ranged aged from 13 to 75 ( $M = 20.45$ ,  $SD = 7.59$ ). The majority of participants ( $n = 328$ ) were university students. The sample characteristics of the sample are presented in Table 1.

### 3.4. Data analysis

Following the design of the test, a descriptive analysis of the items was conducted using SPSS software (version v.26). This entailed the calculation of means, standard deviations, skewness, kurtosis coefficients, and homogeneity item analysis. Subsequently, the psychometric properties were analyzed. Specifically, the validity evidence based on the internal structure was considered through the use of confirmatory factor analysis and measurement invariance, across sex and educational level groups. The internal consistency of the scale was evaluated using the Cronbach alpha coefficient (Celina Oviedo and Campo-Arias, 2005).

To test the three-factor structure proposed by Watson and Glaser (1991) and supported by the Halpern critical thinking tests, a confirmatory factor analysis (CFA) was conducted. To perform this analysis, as recommended for the use with ordinal variables, the unweighted least squares (ULS) estimator, based on the polychoric correlation matrix, was used (Morata-Ramírez et al., 2015). To assess the fit indices of the model, the chi-square statistic ( $\chi^2$ ) and the following indices were calculated: the comparative fit index (CFI), the non-normalized fit index (NNFI), the root mean square error of approximation (RMSEA) with its interval of 90 % interval,

**Table 1**  
Sociodemographic characteristics of the sample ( $N = 575$ ).

Variables	<i>N</i>	%
Sex		
Men	201	35
Women	374	65
Spanish region		
North	210	36.5
South	365	63.5
Type of studies		
Secondary	131	22.8
High school	116	20.2
University discipline		
Social Sciences: Educational Sciences, Psychology, Journalism, Social Education, Librarianship	202	35.1
Philology, Linguistics, Literary Theory, Literature, Translation, Philosophy	52	9.0
Telecommunications, Computing	36	6.3
Health Sciences: Medical Sciences, Nursing, Pharmacy	25	4.3
Law/Business Administration/Marketing/Economics	13	2.3

and the standardized root mean square residual (SRMR). The criteria set forth by Hu and Bentler (1999) were employed to interpret the indices. Values exceeding 0.95 or higher indicate a satisfactory fit for CFI and NNFI; while values between 0.06 and 0.08 signify a reasonable fit for RMSEA (Browne & Cudeck, 1993); and a value approaching 0.08 is deemed a favorable fit for SRMR (Hu & Bentler, 1999).

Accordingly, measurement invariance with respect to sex and educational level have been level has been investigated through the calculation of three nested invariance models, which impose successive restrictions: configural (factor structure similarity within groups), metric (the loading pattern equality), and scalar (intercept similarity). In order to evaluate the fit of our model, we combined the aforementioned criteria. The chi-square test, the ratio of maximum-likelihood chi-square to the degrees of freedom ( $\chi^2/df$ ), the comparative fit index (CFI), and the root mean square error of approximation (RMSEA) were employed for this purpose. In order to ascertain the goodness-of-fit statistic of the level of invariance between the models, the following cut-off points have considered: a change of 0.015 or higher in RMSEA in conjunction with changes in CFI of 0.010 would indicate that there is no invariance. Moreover, in order to comply with the  $\chi^2/df$  ratio, a range of 0 to 3 has been proposed as a threshold for determining the goodness of fit. The Lavaan package (version 0.6–13) in the RStudio terminal was used to analyze the multigroup CFA.

#### 4. Results

Table 2 shows the descriptive results, indicating specifically the means, standard deviations, skewness and kurtosis statistics for each item and dimension. Some skewness and kurtosis values indicate deviation from normality. Also, the item analysis showed values of the corrected item-dimension correlation greater than the cut-off point of 0.30, thus, all items showed reasonable homogeneity.

In particular, regarding the evidence pertaining to the internal structure of the test, the results of the Confirmatory Factor Analysis (CFA) conducted on the entire sample of 575 participants are presented in Table 3. The model with three dimensions demonstrated an excellent fit to the data, as evidenced by values of CFI and NNFI exceeding 0.95. Additionally, the RMSEA levels fell within an acceptable range of 0.06 to 0.08, and the SRMR values approached 0.08, further supporting the model's suitability. Table 3 also presents the results of the configural, metric, and scalar invariance analyses between sexes and between two educational level groups. A decrease in the CFI of <0.01, accompanied by an increase in the RMSEA of <0.015 from the least restricted to the most restricted model, is indicative of a good fit.

In conclusion, the results of the CFA, illustrated in Fig. 1, substantiate the validity of the internal structure of the proposed assessment tool, specifically in terms of construct validity.

To ascertain the reliability of the scores, the Cronbach's alpha coefficient was employed. The coefficients indicated satisfactory reliability of the scores. Specifically, the Cronbach's alpha coefficient yielded values of 0.77 for dimension 1, 0.86 for dimension 2, and 0.84 for dimension 3, respectively.

#### 5. Discussion and conclusions

In this study, a critical thinking test named CRISENSE was developed, and its psychometric features were examined. The test was designed to measure the three dimensions of critical thinking (Observe, Think, Act). The study was conducted in two phases. The initial phase involved focus groups with professionals in the fields of education and psycho-pedagogy. In this phase, a battery of items pertaining to the three identified dimensions was developed, and its items were evaluated by a panel of experts. The aim was to obtain

**Table 2**  
Descriptive statistics for items and dimensions, and corrected item-dimension correlation.

Item/dimension	M	SD	Skewness	Standard error	Kurtosis	Standard error	Corrected item-dimension correlation
Dimension 1	19.88	5.65	-0.51	.102	0.58	.203	
1	4.13	1.21	-1.43	.102	0.97	.203	.63
4	2.70	1.55	0.26	.102	-1.46	.203	.47
7	3.08	1.34	-0.05	.102	-1.10	.203	.54
10	3.45	1.56	-0.40	.102	-1.40	.203	.48
13	4.09	1.24	-1.24	.102	0.40	.203	.58
16	2.44	1.32	0.64	.102	-0.67	.203	.45
Dimension 2	22.15	6.11	-1.15	.102	0.95	.203	
2	3.94	1.20	-1.05	.102	0.17	.203	.77
5	3.77	1.42	-0.76	.102	-0.86	.203	.53
8	3.51	1.29	-0.48	.102	-0.89	.203	.66
11	3.99	1.30	-1.05	.102	-0.18	.203	.68
14	3.74	1.36	-0.81	.102	-0.61	.203	.71
17	3.20	1.37	-0.31	.102	-1.15	.203	.59
Dimension 3	22.37	6.01	-1.17	.102	1.14	.203	
3	3.86	1.29	-0.92	.102	-0.30	.203	.74
6	3.03	1.55	-0.03	.102	-1.50	.203	.45
9	3.53	1.41	-0.52	.102	-1.03	.203	.62
12	4.24	1.21	-1.54	.102	1.16	.203	.74
15	3.95	1.28	-1.21	.102	0.42	.203	.54
18	3.77	1.26	-0.73	.102	-0.56	.203	.71

**Table 3**

Fit indices for confirmatory factor analysis and for measurement invariance, considering sex and educational level.

Model	$\chi^2$	df	CFI	NNFI	RMSEA	RSMR	$\Delta$ CFI	$\Delta$ RMSEA
Total sample (N = 575)	541.012	132	.984	.982	.073 [0.067, 0.080]	.074		
Sex								
Women (n = 375)	336.627	132	.988	.986	.064 [0.056, 0.073]	.073		
Men (n = 202)	378.106	132	.973	.968	.097 [0.085, 0.108]	.105		
Configural invariance	2115.284	264	.970	.965	.156 [0.15, 0.163]	.072		
Metric invariance	2249.263	279	.968	.965	.157 [0.151, 0.163]	.075	<0.01	<0.015
Scalar invariance	2284.571	294	.967	.966	.154 [0.148, 0.160]	.076	<0.01	<0.015
Educational level								
Universitaries (n = 328)	327.193	132	.984	.981	.067 [0.058, 0.076]	.076		
Under universit. (n = 247)	411.834	132	.979	.976	.093 [0.083, 0.103]	.099		
Configural invariance	2163.24	264	.969	.964	.158 [0.152, 0.165]	.074		
Metric invariance	2297.87	279	.967	.964	.159 [0.153, 0.165]	.076	<0.01	<0.015
Scalar invariance	2338.49	294	.967	.965	.156 [0.150, 0.162]	.077	<0.01	<0.015

Note.  $\chi^2$  = Chi-square statistic; df = degrees of freedom; CFI = comparative fit index; NNFI = non-normed fit index; RMSEA = root mean square error of approximation with 90 % confidence interval;  $\Delta$  CFI = CFI less constrained model – CFI more constrained model;  $\Delta$  RMSEA = RMSEA less constrained model – RMSEA more constrained model.

evidence of validity based on the content of the test. The second phase involved an exhaustive psychometric analysis of the resulting test in a sample of 575 participants. This included a descriptive and homogeneity analysis of the items, an evaluation of the internal structure's validity, an investigation of measurement invariance according to sex and educational level, and a reliability analysis.

There has been considerable debate surrounding the role of critical thinking in education as a fundamental aspect of teaching and learning. The objective is to develop critical citizens capable of making decisions without being unduly influenced by external factors such as advertising or ideology. This would facilitate several beneficial outcomes, including the fight against sect recruitment, the improvement of society thought greater empathy (putting we in the shoes of others), objective thinking, and the prevention of fake news from gaining a foothold in communities with well-developed structures of constructive criticism. However, thus far, it has not been considered that thinking should be a de facto tool, or, in other words, a *modus vivendi*. The Spanish education system (at all levels of implementation, from the national to the autonomous, based on the guidelines set by the European Union (2018), espouses the dichotomy of being and doing versus wanting to be and to say.

At present, critical thinking tests are overly reliant on linguistic and cultural context, which is the rationale behind the development of this pioneering tool that assesses logical processes. As no other tool is arranged in this way, it has not been possible to compare the results, as they do not measure the same thing.

The tools used in the field of Spanish are translations from English, which are constrained by the inherent limitations of translating in a technical domain. The nuances and inferences inherent to different cultures make the accurate implementation of these tools challenging, and in some instances, impede comprehension of the intended meaning. Consequently, the necessity arises for developing an *ex professo* instrument that responds to the requirements of our language, while also transcending the conventional notion of measurement and focusing on competence. In this regard, even the dimensions have been translated, with the referents exhibiting that renders it exceedingly challenging to delineate reality in a manner that is identical in English as it is in Spanish. As is the case with the distinction between “ability” and “competence” or “intelligence” in English the concepts employed by Halpern, Watson, and Glaser, cannot be translated literally. The proposed tool was designed to be applied in such contexts. It was implemented in two phases (by a panel of experts and by students), and its results were analyzed for validation with positive results.

Considering the skills required of citizens by governmental bodies, both in the context of pedagogy and social cooperation, the initial phase in progressing appropriately is the construction of a conceptual model for critical capacity that surpasses mere critical thought. This model must also delineate the facets of critical capacity in a precise and succinct way. This proposition provides a framework for the evaluation of critical expertise from adolescence onwards in both academic and vocational settings.

In accordance with the guidelines set forth by AERA, APA and NCME (2014), CFA was identified as the optimal method for evaluating the construct validity of the instrument. This approach was deemed the most suitable for providing evidence of the internal structure of the test and for substantiating the model's prior theoretical establishment. In other words, the latent factors considered in the conducted CFA were aligned with the three dimensions that had been established for the assessment instrument.

In examining the psychometric properties of the test, the evidence supporting the internal structure's validity has been evaluated, as well as the invariance of measurements according to sex, the analysis of items, and the reliability of test scores. In conclusion, the test exhibits satisfactory psychometric properties, as evidenced by the support of the three-dimensional model with favorable fit indices. Regarding the measurement invariance analysis, it can be concluded that the measure is suitable for both men and women, as well as for two distinct groups of students, in that it allows for the measurement of the studied construct in a consistent manner. Moreover, the test exhibits satisfactory reliability, as evidenced by alpha coefficients exceeding 0.70.

It is important to acknowledge the limitations of the present study. Firstly, it should be noted that a convenience sample was used, which may limit the generalizability of the results. It would be beneficial for future research to conduct an experimental investigation with a representative sample of the target population, as is already done with other tests such as PISA or PIRLS. Secondly, the fit indices for the measurement invariance analysis could be enhanced by increasing the sample size, as the division of the sample into the corresponding groups has an impact on the statistical power of the analysis. Accordingly, future research will expand the sample size

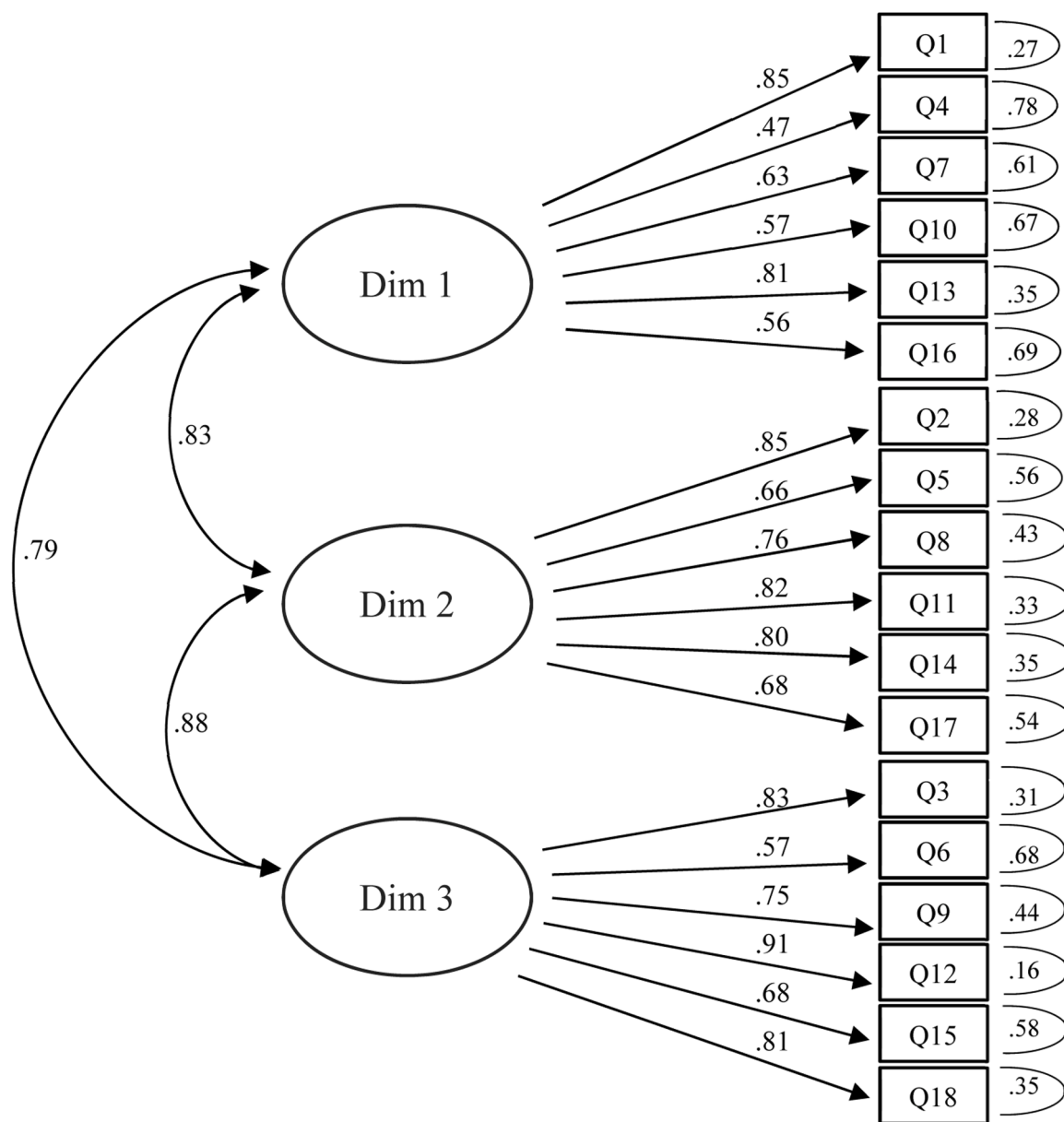


Fig. 1. Standardized factor loadings for the test in the total sample.

by replicating the analyses and expanding the psychometric information on the subject. Notwithstanding these limitations, the study also possesses notable strengths.

This study establishes a foundation for future studies with a larger sample that is representative of the intended sample: students from all educational stages of the Spanish system, as well as any adult who uses Spanish as a vehicular language. Moreover, it can be employed in an academic context to illustrate the critical competence skills as delineated by the European Union. Furthermore, this tool can be utilized in a variety of professional contexts where critical competence is a valuable asset, streamlining the recruitment process for human resources professionals. The dimensions were developed from a theoretical perspective with the intention of enabling both the measurement of the construct and the training thereof. This constitutes the initial stage in the scientific validation of the content. The second step will be to implement the knowledge transfer from the university to the classroom setting through the administration of OTA-based interventions and the subsequent measurement of student progress. The objective is to provide a valuable instrument in the classroom that enhances the teaching-learning process, rather than merely assessing individual competence without a practical objective.

In any case, this test and its dimensions have been designed to be applicable in any language, as they represent situations that require observation, critical thinking, and decision-making. It should be noted that the linguistic component may still exert an



influence despite the limitations. A further objective would be to study this relationship and quantify it. Similarly, it would be beneficial to delineate and measure the influence of experience (age) on this construct, to gain insight into the relationship between the two. Consequently, as this study employs a novel approach, there are no existing studies with which to compare the results. Although this approach may offer an innovative and disruptive perspective on today's scientific literature, it is important to note that taking a first step can involve the errors of not having similar starting points upon which to build.

### CRedit authorship contribution statement

**Elena del Pilar Jiménez Pérez:** Writing – original draft, Investigation, Conceptualization. **Manuel León Urrutia:** Writing – review & editing, Writing – original draft, Validation, Methodology, Investigation, Formal analysis, Data curation. **Pedro García Guirao:** Supervision, Investigation, Conceptualization. **María Victoria Cerezo Guzmán:** Validation, Methodology, Formal analysis, Data curation.

### Acknowledgments

This study would not have been possible without the generous sponsorship of the *Spanish Association of Reading Comprehension* (AECL) and the research group Hum-1048 (University of Malaga), of which several members are affiliated. The research visits of all members to other institutions, in which they have participated as visitors from one university to another, have also been instrumental in advancing the project. Furthermore, the research project is part of the GIE Innova22 (UMA, PIE22–111).

### Data availability

Data is available in the Pure's University of Southampton repository at <https://doi.org/10.5258/SOTON/D3296>

### References

- Al-Ahmadi, F. M. A. (2008). The development of scientific thinking with senior school physics students. Is construct validation valid? *Philosophy of Science*, 83(5), 1098–1109. <https://doi.org/10.1086/687941>
- American Educational Research Association, American Psychological Association, and National Council for Measurement in Education (AERA, APA, and NCME). (2014). *Standards for educational and psychological testing*.
- Bachelard, G. (1994). *La formación del espíritu científico*. Siglo XXI Editores.
- Baghana, J., Bondarenko, E. V., Blazhevich, Y. S., Zubritskij, O. A., & Slobodová Nováková, K. (2020). The importance of cultural seme in the semantics of different language units for political texts interpretation. *Applied Linguistics Research Journal*, 4(9). <https://doi.org/10.14744/alrj.2020.09581>
- y Betancourth, S., Muñoz, K. T., & Rosas, T. J. (2017). Evaluación del pensamiento crítico en estudiantes de educación superior de la región de Atacama-Chile. *Prospectiva*, (23), 199–223. <https://doi.org/10.25100/prts.v0i23.4594>.
- Blakemore, S. (2018). *Inventing ourselves: The secret life of the teenage brain*. Transworld Publishers.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. *Testing structural equation models* (pp. 136–162). Sage. eds KA Bollen, JS Long.
- Caravita, S., & Hallden, O. (1994). Re-framing the problem of conceptual change. *Learning and Instruction*, 4, 89–111.
- Caro Valverde, M. T., De Vicente-Yagüe Jara, M. I., & Valverde González, M. T. (2018). Percepción docente sobre costumbres metodológicas de argumentación informal en el comentario de texto. *Revista Española de Pedagogía*, 76(270), 273–293. <https://doi.org/10.22550/REP76-2-2018-04>
- Celina Oviedo, H., & Campo-Arias, Adalberto (2005). Aproximación al uso del coeficiente alfa de Cronbach. *Revista Colombiana de Psiquiatría*, 34(4), 572–580. <https://doi.org/10.7705/biomedica.v26i4.327>
- Comisión Europea [European Commission] (2021). [https://ec.europa.eu/education/policies/higher-education/bologna-process-and-european-higher-education-area\\_es](https://ec.europa.eu/education/policies/higher-education/bologna-process-and-european-higher-education-area_es).
- Facione, P. A. (2007). Pensamiento crítico. Qué es y por qué es importante. *EduTEKA*, 23–56.
- Fernández, G., García-Guirao, P., & López, O. (2021). La lectura del Lazarillo de Tormes para estimular el pensamiento crítico en secundaria. *Investigaciones Sobre Lectura*, 16, 1–15.
- García Medina, M. A., Acosta Meza, D., Atencia Andrade, A., & Rodríguez Sandoval, M. (2020). Identificación del pensamiento crítico en estudiantes universitarios de segundo semestre de la Corporación Universitaria del Caribe (CECAR). *Revista Electrónica Interuniversitaria de Formación del Profesorado*, 23(3), 133–147. <https://doi.org/10.6018/reifop.435831>
- Gardner, H. (2005). *Las cinco mentes del futuro*. Paidós.
- González Zamora, J. H. (2006). *Discernimiento: Evolución del pensamiento crítico en la educación superior*. Universidad Icesi.
- Gravet, V., & Hainselin, M. (2021). Book Review: Critical Thinking: A Concise Guide. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.656331>
- Gutiérrez Fresneda, R. (2022). Influencia de las estrategias cognitivas de la lectura en la mejora de la capacidad de comprensión en estudiantes de Educación Primaria. *Investigaciones Sobre Lectura*, 2(17). <https://doi.org/10.24310/isl.vi18.15140>
- Gutiérrez-Fresneda, R., Díez, A., & Jiménez-Pérez, E. (2017). Longitudinal study on learning to read in early ages. *Revista de Educación*, 378, 28–49. <https://doi.org/10.4438/1988-592X-RE-2017-378-360>
- Halpern, D.F. (2012). *Halpern critical thinking assessment: Test manual*. Schuhfried GmbH, Ed.). <https://doi.org/10.5430/wjel.v12n3p11>.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Izarra, D., Hirsch, A., & Rodríguez, A. M. (2020). Profesorado de posgrado y el desarrollo del pensamiento crítico. *Innovación educativa*, 83(20), 11–34.
- Jiménez-Pérez, E., León-Urrutia, M., Rojo, M. V., & García-Guirao, P.: (2024). Design and validation of CRISENSE, a novel critical competence assessment tool for Spanish adolescents and young adults. *Pure*. <https://doi.org/10.5258/SOTON/D3296>
- Jiménez Pérez. (2023). Pensamiento crítico VS competencia crítica: Lectura crítica. *Investigaciones sobre lectura*, 18(1), 1–26. <https://doi.org/10.24310/isl.vi18.15839>.
- Jiménez-Pérez, E., & del, P. (2017). Emotional intelligence as a predictor of reading habits and reading competence in university students. *Investigaciones Sobre Lectura*, 10, 42–54. <https://doi.org/10.24310/revistaisl.vi10.11000>
- Jiménez-Pérez, E., del, P., de Vicente-Yagüe, Jara, M.-I., Gutiérrez-Fresneda, R., & García-Guirao, P. (2021). Sustainable education, emotional intelligence and mother-child reading competencies within multiple mediation models. *Sustainability*, 13, 1803. <https://doi.org/10.3390/sul3041803>
- Kurland, D.J. (2000). Critical Reading VS Critical Thinking. [http://www.criticalreading.com/critical\\_reading\\_thinking.htm](http://www.criticalreading.com/critical_reading_thinking.htm).
- LOE (2006). Ley Orgánica 2/2006, de 3 de mayo, de Educación. [Organic Law 2/2006 on Education of 3 May].

- Marcano Medina, J. E. (2023). *Desarrollando el pensamiento crítico: Superando los prejuicios*. Universidad de Puerto Rico. <http://www.uprh.edu/ccc/ciencias>.
- Mockus, A. (1989). Formación básica y actitud científica. *Educación y Cultura*, 17, 11–16. <http://www.fecoderevistaeducacionycultura.com/>.
- Morata-Ramírez, M. A., Holgado-Tello, F. P., Barbero-García, I., & Méndez, G. (2015). Confirmatory factor analysis. Recommendations for unweighted least squares method related to Chi-Square and RMSEA. *Acción Psicológica*, 12(1). <https://doi.org/10.5944/ap.12.1.14362>, 79-9.
- Mullis, I. V. S., & Martin, M. O. (2019). *PIRLS 2021 assessment frameworks. international study center tims & pirls*. Boston College. [http://pirls2021.org/frameworks/wp-content/uploads/sites/2/2019/04/P21\\_FW\\_Ch1\\_Assessment.pdf](http://pirls2021.org/frameworks/wp-content/uploads/sites/2/2019/04/P21_FW_Ch1_Assessment.pdf).
- Muñiz, J., & Fonseca-Pedrero, E. (2019). Ten steps for test development. *Psicothema*, 31, 7–16. <https://doi.org/10.7334/psico-thema2018.291>
- Nieto, A. M., Sáiz, C., & Orgaz, B. (2009). Análisis de las propiedades psicométricas de la versión española del HCTAES-Test de Halpern para la evaluación del pensamiento crítico mediante situaciones cotidianas. *Revista Electrónica de Metodología Aplicada*, 1(14), 1–15.
- OCDE. (2018). *Marco teórico de lectura*. PISA, 2018 [https://www.educacionyfp.gob.es/dam/jcr:2f1081a1-c1e4-4799-8a49-9bc589724ca4/marco%20teorico%20lectura%202018\\_esp\\_ESP.pdf](https://www.educacionyfp.gob.es/dam/jcr:2f1081a1-c1e4-4799-8a49-9bc589724ca4/marco%20teorico%20lectura%202018_esp_ESP.pdf).
- Palma Luengo, M., Ossa Cornejo, C., Ahumada Gutiérrez, H., Moreno Osorio, L., & Miranda Jaña, C. (2021). Adaptación y validación del test Tareas de Pensamiento Crítico en estudiantes universitarios. *Revista de Estudios y Experiencias en Educación REXE*, 20(42), 199–212.
- Paul, R., & Elder, L. (2003). La mini-guía para el pensamiento crítico. *Conceptos y herramientas. fundación para el pensamiento crítico*. <http://www.criticalthinking.org>.
- Puche, R. (2000). *Formación de herramientas científicas en el niño pequeño*. Universidad del Valle - Arango editores.
- Redondo Gutiérrez, L., Corrás, T., Vázquez, M. J., & Fariña, F. (2017). La capacidad crítica como competencia transversal: La evaluación como herramienta. *Revista de Estudios e Investigación en Psicología y Educación*, (07), 097–101. <https://doi.org/10.17979/reipe.2017.0.07.2734>
- Robinson, K., & Aronica, L. (2009). *El elemento*. Grijalbo.
- Rodríguez-Gómez, G., Ibarra-Sáiz, M. S., Gallego-Noche, B., Gómez-Ruiz, M.Á., & Serra, V. Q. (2012). La voz del estudiante en la evaluación del aprendizaje: Un camino por recorrer en la universidad. *RELIEVE. Revista Electrónica de Investigación y Evaluación Educativa*, 18(2), 1–21. <https://doi.org/10.7203/relieve.18.2.1985>
- Seligman, M. E. P. (2005). *La auténtica felicidad*. Byblos.
- Sharma, M., Doshi, B. M., Verma, M., & Verma, A. K. (2022). Strategies for Developing Critical-Thinking Capabilities. *World Journal of English Language*, 12(3), 117.
- y Spelke, E. (1991). Physical knowledge in infancy: Reflections on Piaget's theory. In S. Carey, & R. Gelman (Eds.), *The epigenesis of mind: Essays on biology and cognition*. Erlbaum.
- Tamayo, O. E. (2009). *Didáctica de las ciencias: La evolución conceptual en la enseñanza y el aprendizaje*. Centro Editorial Universidad de Caldas.
- Tamayo, O. E., Zona, R., & Loaiza, Z. Y. (2014). *Pensamiento crítico en el aula de ciencias*. Universidad de Caldas.
- UNESCO. (1995). *Documento de política para el cambio y el desarrollo en la educación superior*. UNESCO. <http://unesdoc.unesco.org/images/0009/000989/098992s.pdf>.
- UNESCO. (1998). *Conferencia mundial sobre la educación superior la educación superior para el siglo XXI: Visión y acción*. UNESCO. <http://unesdoc.unesco.org/images/0011/001163/116345s.pdf>.
- UNESCO. (2009). *Comunicado conferencia mundial sobre la educación superior: Las nuevas dinámicas de la educación superior y la investigación para el cambio social y el desarrollo*. UNESCO. <http://www.scielo.org.mx/pdf/peredu/v31n126/v31n126a8.pdf>.
- Vicente-Yagüe Jara, M. I., Valverde González, M. T., & González García, M. (2019). Necesidades de formación del profesorado de Lengua y Literatura para el desarrollo de la argumentación informal en el comentario de texto. *Educación Siglo XXI*, 37(1), 213–234. <https://doi.org/10.6018/educatio.363471>
- Vincent-Lancrin, S., González-Sancho, C., Bouckaert, M., Luca, F., Fernández-Barrera, M., Jacotin, G., Urgel, J &, & Vidal, Q. (2019). *Fostering students' creativity and critical thinking. what it means in school*. OECD Publishing. <https://doi.org/10.1787/62212c37-en>
- Watson, G., & Glaser, E. M. (1991). *Watson-Glaser critical thinking appraisal manual*. London: Psychological Corporation.