Migration Letters

Volume: 20, No: S9(2023), pp. 1666-1692

ISSN: 1741-8984 (Print) ISSN: 1741-8992 (Online)

www.migrationletters.com

PCB Test to Assess Critical Thinking in Students from Disadvantaged Contexts

García Medina María Angélica[1], Acosta Meza David de Jesús[2], Lengua-Cantero Claudia[3*], Anaya Herrera Jhon Édison[4]

Abstract

The study focused on the design of a test to evaluate the critical thinking of university students in the Colombian Caribbean Region, based on the results of the Saber 11 standardized tests, applied by the Colombian Ministry of National Education, to measure the basic competencies of eleventh-grade students, among which is critical reading. The research was developed in three phases: review and theoretical referencing of the different instruments for measuring critical thinking, design of the dimensions and indicators, according to the disadvantaged contexts, and validation of the instrument. The research approach is mixed. The statistical instruments CVI, KMO, exploratory factor analysis, confirmatory factor analysis, and Barlett's test of sphericity reveal reliability in the qualitative categories of the test, confirmatory results in the dimensions and indicators, and coherence among the items. When the pilot test was applied, the conceptual and contextual relationship of the dimensions and indicators was found. The study suggests that the PC-B test is an effective tool for gauging the critical thinking abilities of students from disadvantaged backgrounds.

Keywords: *Context, thinking skills, critical thinking, test, validation.*

1. Introduction

A prominent difficulty faced by students entering university studies is the reading and comprehension levels at the end of secondary education, according to the results of the standardized test “Saber 11”, evaluated by the Ministry of National Education (MEN). In addition, in 2000, the Organization for Economic Cooperation and Development (OECD) test, the Program for International Student Assessment (PISA), was applied, focused on a young population between 15 and 17 years of age. The performance of Colombia in these tests was an average between 3.85 and 4.12, slightly lower than in previous years.

It is necessary to contrast the contextual learning needs and the socio-demographic situations of the students who apply to the university programs of the Corporación Universitaria del Caribe CECAR; they come from socioeconomic strata 1, 2 and 3, according to the National Administrative Department of Statistics (DANE), which

[1] Faculty of Basic Sciences, Engineering and Architecture, Corporación Universitaria del Caribe CECAR, Western Troncal Highway, Km. 1, Vía Corozal - Sincelejo, Colombia, maria.garciame@cecar.edu.co

[2] Faculty of Humanities and Education, Corporación Universitaria del Caribe CECAR, Western Troncal Highway, Km. 1, Vía Corozal - Sincelejo, Colombia, david.acosta@cecar.edu.co

[3] Faculty of Basic Sciences, Engineering and Architecture, Corporación Universitaria del Caribe CECAR, Western Troncal Highway, Km. 1, Vía Corozal - Sincelejo, Colombia, claudia.lengua@cecar.edu.co

[4] Faculty of Humanities and Education, Corporación Universitaria del Caribe CECAR, Western Troncal Highway, Km. 1, Vía Corozal - Sincelejo, Colombia, jhon.anaya@cecar.edu.co

categorizes the population based on unsatisfied basic needs, health, sanitation, education, and essential services.

The objectives of the project are based on the fundamental, theoretical, and contextual needs that will make viable the scope of the research purposes and the object of study; it was proposed to design the PC-B test, to determine the levels of development of critical thinking in first-year university students, from three phases in the research process, to identify the conceptual and methodological references, to characterize the levels of development of critical thinking. To design the test, several bibliographic reviews were carried out, contrasting through a referencing matrix the tests published in the academic communities; based on this review, the dimensions and indicators were proposed according to the needs of the context.

The critical thinking test, aimed at assessing foundational skills, proposed to be applied to first-year students in the aforementioned academic programs comprises the following dimensions: observation, comparison, classification, and description, subdivided into the indicators: identifying, reporting, detecting, defining, relating, differentiating, proposing, characterizing, associating, analyzing, listing and reflecting, as illustrated in Figure 1. Likewise, this study emphasized testing how these are related and the dependence between them through a structural equation model (Escobedo et al., 2016), findings that evidence the reliability and feasibility of the results.

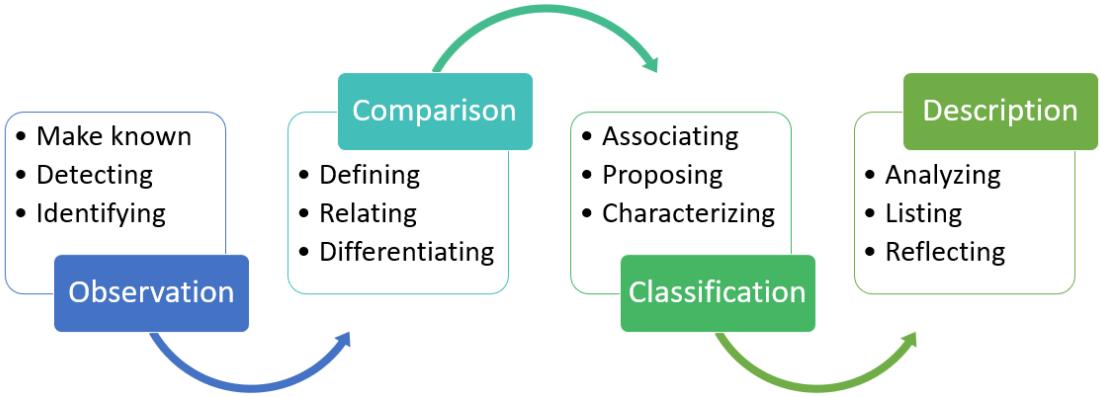

Figure 1. Dimensions and indicators of the del Critical Thinking Test – PC-B Test.

1.1 Critical thinking in higher education

Critical thinking is a process of seeking knowledge through reasoning skills, problem-solving, and decision-making that allows achieving the desired results with the greatest efficiency (Saiz and Rivas, 2011). This concept corresponds to that of Halpern (2014) who conceives critical thinking as the search for collecting information from the process of thinking and inquiring knowledge, together with its fundamental mechanisms, aimed at achieving the established purposes. In addition, Guevara (2000) points out that they are those thinking skills that serve to survive in the everyday world, have a social function, and seen in this way, the student must not put them aside.

Critical thinking is rational, reflexive, and interested in what to do or believe. It is the mental activity that allows evaluating the formulated arguments or propositions that guide the development of beliefs and the taking of action (Sánchez, 2013). It is the ability to analyze facts, generate and organize ideas, defend opinions, make comparisons, make inferences, evaluate arguments, and solve problems. It is a conscious and deliberate process for interpreting or evaluating information and experiences using attitudes and skills that guide fundamental beliefs and actions (Sanchez, 2013).

Critical thinking is the intellectually disciplined process of conceptualizing, applying, analyzing, synthesizing, or evaluating data generated through observation, experience,

reflection, reasoning, or communication to guide belief and action (Scriven and Paul, 2003).

In this study, critical thinking represents the liberating force enabling individuals to question, develop curiosity, reflect, analyze, and investigate. It is also the mental process where individuals prioritize reason, flexibility, and impartiality when judging environmental problems.

In recent years, Higher Education Institutions have implicitly incorporated in their curricula the development of critical thinking as a soft competence. However, some associated factors such as cultural, socioeconomic, and cognitive factors pose challenges to the achievement of this (Gyenes, 2021).

At the national level, strategies have been developed to promote critical and reflective thinking in students to build a student capable of adapting to any context with a global vision of thinking to transform this society in the crisis of values and social responsibility.

In the regional context, initiatives have been developed within the framework of critical thinking from high school to university level, aiming to foster students who can reflect and propose alternative solutions to contextual problems. Similarly, the curricula emphasize learning outcomes that shape an ethical professional.

The learning outcomes are shared across the different professions. They are related to the implementation in an integral way to the aptitudes, personality traits, and appropriate knowledge. The appropriation of knowledge is understood as a dynamic process that requires a series of systematized actions, integrated and parallel to the university curriculum, in function of high performance (Polo et al., 2022, p. 161).

1.1.1 Critical thinking assessment

The assessment of thinking is a complex task, as it involves establishing a series of concepts, dimensions, and indicators, in many cases abstract. It acts as a diagnostic and prognostic tool on the levels of critical thinking of students to recognize the learning processes in the formation of analytical thinkers from solid and relevant methodologies, taking into account the diversity of capabilities and potentialities of people (Saiz and Rivas, 2008; Alquichire and Arrieta, 2018). Skills refer to the abilities that allow evaluating information, identifying assumptions, deducing, reasoning, understanding, and thinking the logic of concepts to make judgments (Norris and Ennis, 1989).

To measure critical thinking there are qualitative and quantitative instruments designed to be applied in large and small samples (Ossa-Cornejo et al., 2017). In the present study, 14 instruments with their respective authors and subscales are referenced. Although these differ in methodological approaches and subscales, it is also observed that some of these present in common the following dimensions: analysis, evaluation, decision-making, writing, and problem-solving. These instruments are contrasted below in Table 1, and they were the basis for determining the dimensions that would be part of the PC-B Test.

It's crucial to recognize that evaluating students' critical thinking, using concrete knowledge and decision-making modeled on real-life situations, can facilitate feedback through an evidence-based model (Hyytinen et al., 2021). It allows for ongoing examination of thought processes and ensures qualitative and quantitative validity.

2. Methodology

The study was conducted under the mixed research paradigm, which is that in which "quantitative and qualitative research are intertwined, through the use of numerical, verbal, textual, visual, symbolic and other types of evidence to understand problems of the context in which the subjects of study will develop" (Hernández-Sampieri and Mendoza, 2018, p. 10). It was also based on a non-experimental design, understood as

one in which "already existing situations are observed without the deliberate manipulation of variables, only the phenomena are observed in their natural environment and then the results are analyzed", (Hernández-Sampieri and Mendoza, 2018, p. 175). According to the above, the established dimensions and indicators were analyzed without requiring the intervention or manipulation of these to measure the level of critical thinking achieved by a student. In addition, it had an explanatory scope, given that correlations are established between the dimensions and indicators that allow determining the possible causes that lead to the development of critical thinking.

The population that took part in the piloting of the PC-B Test, which measures basic critical thinking skills, consisted of 259 first-year students (those in their first and second semesters) belonging to the following programs: Systems Engineering (31. 7%), Industrial Engineering (13.1%), Social Work (17.8%), Psychology (0.39%), Law (8.88%), Accountancy (2.32%), Sports Science (0.39%), Architecture (12.0%) and Business Administration (13.5%).

The test participants meet the following sociodemographic conditions: The average age is 20 years; the minimum is 16, and the maximum is 29. The latter is considered overage for a first-time student. Regarding gender, 43.2% are female, 56.4% are male, and 0.4% do not indicate gender.

As for the preceding level, 17.8% of the students graduate from rural schools and 82.2% from urban areas. Of these, 24.3% graduate from private sector schools, while 77.7% graduate from public sector schools.

Regarding the residence of the students, 94.2% reside in the department of Sucre, 5.41% in the department of Córdoba, and 0.39% in the department of Bolívar. The graph shows the municipalities where the students live.

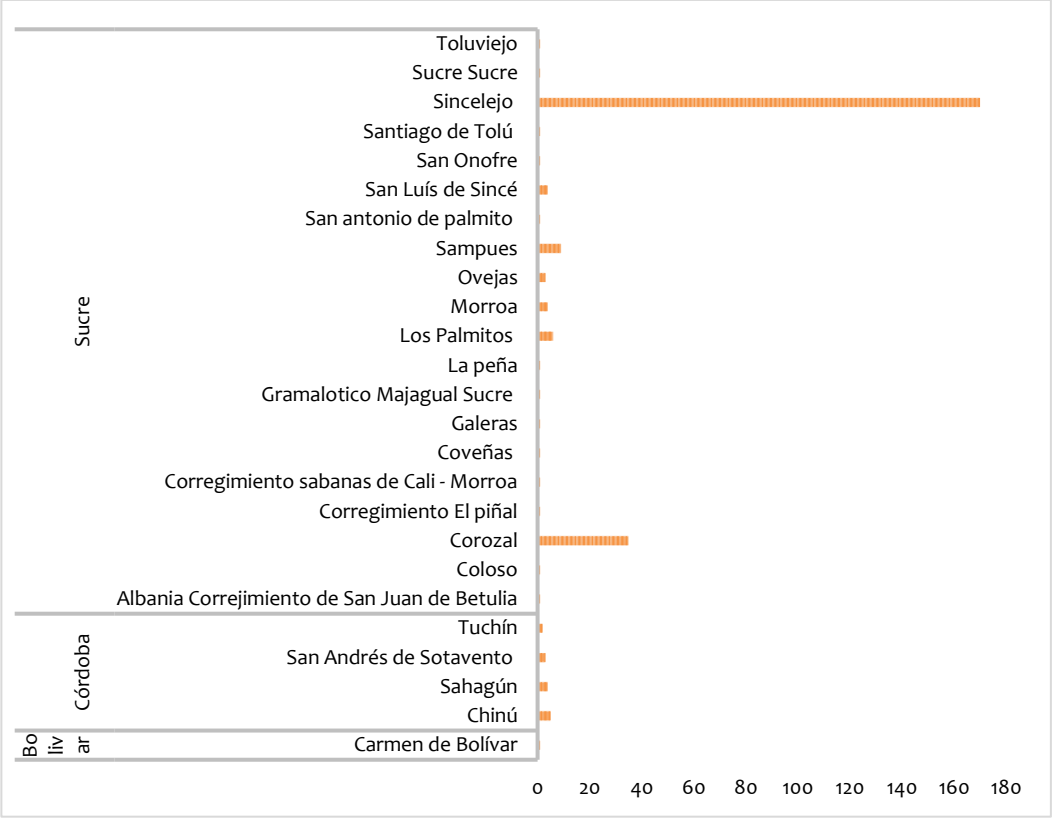

Figure 2 Current residence of students.

The above data allow us to have a global vision of the social and cultural context of the students, influenced by the characteristics of each municipality.

2.1 Field work and data analysis

The Basic Critical Thinking Test - PC-B Test - originated as a need to identify the thinking skills of incoming students, according to the social and economic context. From this need, the objective of the study was to design a test to assess the basic level of critical thinking of first-year university students.

Instruments applied in different contexts, which were proposed by both classical and contemporary authors. The information was organized in a matrix in which the methodological approaches and subscales of these instruments were contrasted (see Table 1). From the documentary review, the following authors are listed: Watson and Glaser, 1981; Ennis, Millman, and Tomko, 1985; Ennis and Weir, 1985; Facione, 1991; Facione, Facione and Sanchez, 1994; Kamin, O'Sullivan, Younger and Deterding, 2001; Watson and Glaser, 2002; Ennis, Millman and Tomko, 2005a; Ennis, Millman, and Tomko, 2005b; Halpern, 2006; Valenzuela and Nieto, 2008; Valenzuela, Nieto, and Saiz, 2011; Ennis, Millman, and Tomko, 2012; Rivas and Saiz, 2012; Calle, 2014; Carrasco, 2018; and Paul and Elder, 2019.

In the second phase, considering the theoretical review, the measurement to evaluate critical thinking was structured. In other words, the operational definitions underlying the test were formulated, as well as the dimensions, indicators, and items that comprise it; the design of the PC-B test was carried out at the beginning of 2021.

In the third phase, the PC-B test was submitted to the validation of academic peers to contrast it with the Content Validation Index (CVI) test established by Tristán-López (2008). The observations made by these experts were used to make adjustments to the response items of the contexts performed. To determine the feasibility of the test, validation was carried out through the KMO test (Kaiser-Meyer-Olkin), which made it possible to establish the congruence between items, indicators, and dimensions. Similarly, Barlett's test of sphericity was applied for hypothesis testing. The confirmatory factor analysis was then applied to corroborate the correlation between variables.

3. Results

3.1 PC-B Test Structuring

Based on the documentary review of the instruments used to measure critical thinking, the authors designed the Test of Critical Thinking at a basic level, called the Test PC-B.

The methodology used to carry out this review consisted of selecting the instruments used frequently to identify common or different elements among them. Then, a conceptual comparison matrix was elaborated in which the name of the instruments (tests) and their authors, and structural aspects such as methodological approach and subscales were recorded (see Table 1).

Table 1 Matrix for contrasting critical thinking measurement instruments.

Test	Authors	Methodological approach	Subscale
California Critical Thinking Skills Test (CCTST)	Facione, P., 1991	Quantitative	- Analysis - Evaluation - Inference - Deductive reasoning - Inductive reasoning

California Critical Thinking Disposition Inventory (CCTDI)	Facione, N., Facione, P. and Sánchez, 1994	Quantitative	- Search for truth - Open-mindedness - Willingness to analyze - Willingness to systematize - Self-confidence in reasoning - Curiosity - Maturity of judgment
Kamin Method	Kamin, O'Sullivan, Younger and Deterding, 2001	Qualitative	- Identification of the problem - Problem description - Examination of the problem - Applicability - Integration - Critical evaluation
Cornell Critical Thinking Test (CCTT)	Ennis, Millman and Tomko, 1985; Ennis, Millman and Tomko, 2005a	Quantitative	- Deduction - Induction - Presumption - Observation/credibility - Meaning
Watson-Glaser Critical Thinking Appraisal (WGCTA)	Watson y Glaser, 1981; Watson and Glaser, 2002	Quantitative	- Inference - Recognition - Interpretation - Argument evaluation
Cornell Critical Thinking Tests, Level X	Ennis, Millman y Tomko, 1985; Ennis, Millman and Tomko, 2005b	Quantitative	- Induction - Deduction - Credibility - Identification of assumptions
Cornell Critical Thinking Tests, Level Z	Ennis, Millman and Tomko, 2012	Quantitative	- Induction - Deduction - Credibility - Identification of assumptions

			- Semantics - Definition - Prediction in planning experiments
The Ennis-Weir Critical Thinking Essay Test	Ennis and Weir, 1985	Qualitative	- Information - Conclusions - Assumptions - Concepts - Transcendence - Point of view
HCTAES – Halpern Critical Thinking Assessment Using Everyday Situations	Halpern, 2006	Mixed	- Hypothesis testing - Verbal reasoning - Argument analysis - Probability and uncertainty - Decision making - Problem solving
Scale of observation of critical thinking skills associated with writing	Calle, 2014	Quantitative	- Establishment of a position on an issue - Construction of arguments - Drawing conclusions - Self-regulation
Motivational Scale of Critical Thinking (EMPC, as its acronym in Spanish)	Valenzuela and Nieto, 2008; Valenzuela, Nieto y Saiz, 2011	Quantitative	- Expectancy - Task Value o Importance o Utility o Interest o Cost
Contextualized Critical Thinking Evaluation EP-2C	Carrasco, 2018	Qualitative	- Analysis - Interpretation - Evaluation - Inference - Explanation
Critical thinking test - PENCRISAL	Rivas y Saiz, 2012	Qualitative	- Deductive reasoning - Inductive reasoning - Practical reasoning - Decision making

		- Problem solving
The International Critical Thinking Reading and Writing Test	Paul and Elder, 2019	Mixed	- Thinking analysis - Assessment of thinking - Thinking dispositions - Thinking skills and abilities - Obstacles or barriers to critical thinking

After examining the methodologies and subscales of the tests that measure critical thinking in an individual, a cross-checking of the subscales was carried out, which made it possible to establish the dimensions and indicators that make up the PC-B Test (see Figure 3).

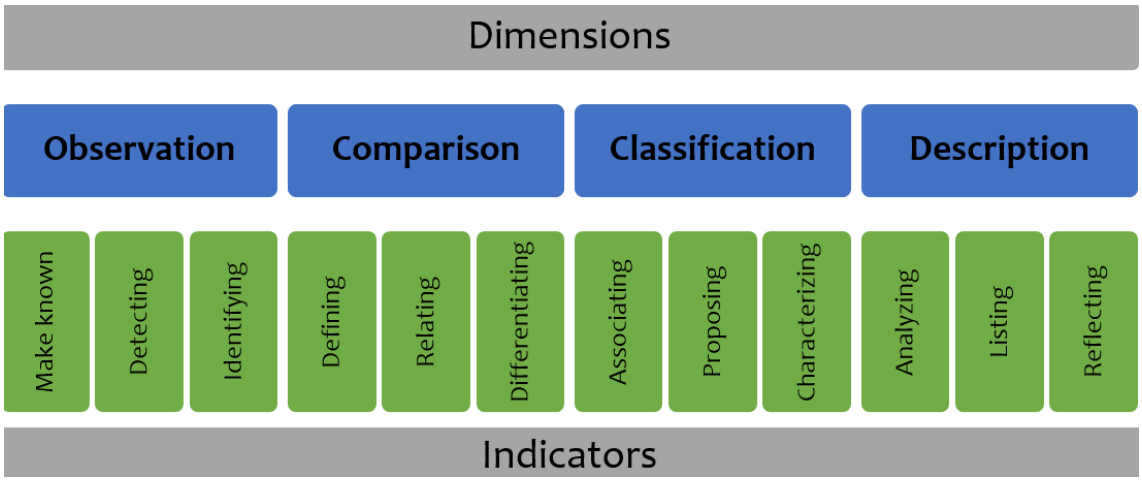

Figure 3. Dimensions and indicators PC-B Test

The PC-B Test makes it possible to determine the ability of the students to apply basic critical thinking skills accurately following their academic background. These skills derive from four dimensions that provide reliable and valid results. The test items derive from on multiple-choice and multiple-choice questions with a single answer, with contextualized themes, situations, and problems that students face daily. Likewise, these have been structured based on colloquial language, which facilitates the understanding of the statements by students in the first years of higher education. The PC-B test consists of 36 questions, broken down as follows: 3 questions for each indicator and 9 for each dimension, as illustrated in Figure 4.

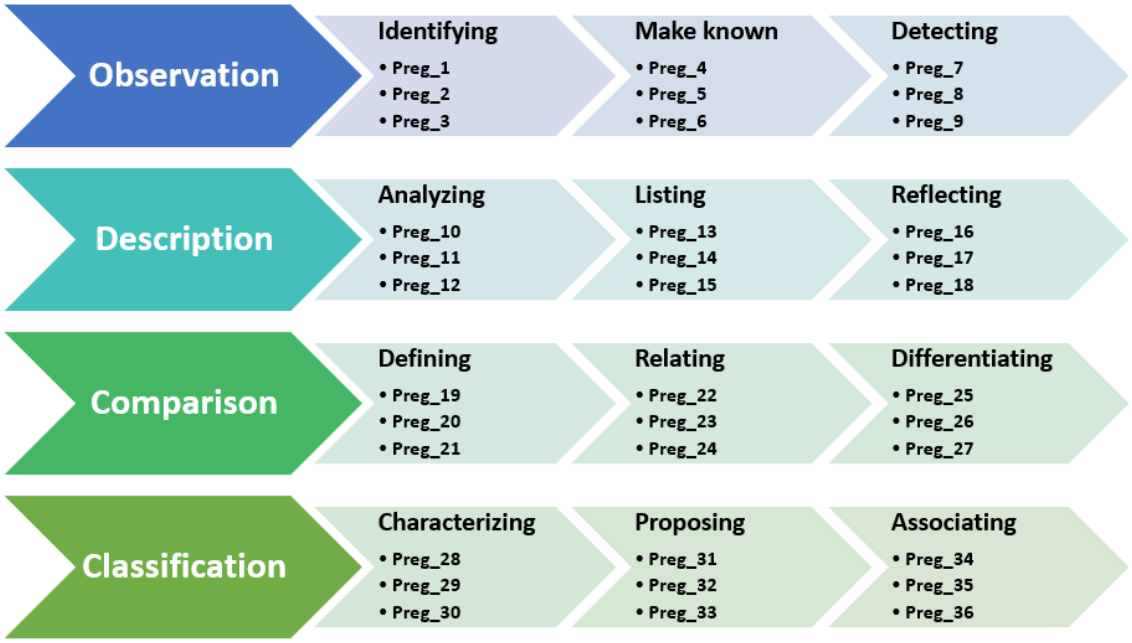

Figure 4 Order of questions according to dimension and indicators

The dimensions and indicators of the test are described below, taking into account the review of bibliographic sources on critical thinking.

Observation: the mental process that allows identifying and reporting the characteristics of the observed objects to distinguish conceptual representations at a given moment (De Sánchez, 1995; Velásquez et al., 2013). The indicators associated with this dimension are: a) identify: the ability to observe the characteristics of an object to be recognized in the environment (Bigg, 2006; Arredondo, 2014); b) reporting: provide coherent information about what is known or understood about an object or subject (García et al., 2019) and c) detect: perception of a given fact that is manifested through physical methods (Soca, 2007).

Comparison: Refers to the extension of observation to establish identifications and relationships of objects through mental representations that allow the abstraction of differences (De Sánchez, 1995; Velásquez et al., 2013). The indicators associated with this dimension are: a) defining: understanding the observed problems to list and identify their characteristics (Márquez, 2014; Roca, 2013), b) relating: process in which concepts are associated and applied to problems in context (De Sánchez, 1995; Marzano, 2001; Bigg, 2006; Márquez, 2014), and c) differentiating: distinguishing dissimilar and similar characteristics of objects (Arredondo, 2014).

Classification: It is considered the mental process by which an individual can identify and differentiate similarities between observed objects to understand and apply concepts through mental representations (De Sánchez, 1995; Velásquez et al., 2013). The indicators associated with this dimension are: a) characterizing: establishing descriptive particularities of objects to define qualities or attributes (Strauss and Corbin, 2002; Sánchez, 2010; Pérez and Gardey, 2017), b) proposing: expressing thoughts about an object in a propositional way by grouping them into criteria and concepts (Pérez and Merino, 2013) and c) associating: establishing mental relationships between objects involving conceptual connections (Arredondo, 2014).

Description: reporting the characteristics of the object in a precise and orderly manner through deductive and inductive thinking (De Sánchez, 1995; Velásquez et al., 2013). The indicators associated with this dimension are a) listing – the process of recording and noting objects in an orderly manner with a fixed purpose (De Sánchez, 1995; Bigg, 2006; Velásquez et al., 2013); b) analyzing – intellectual action with which objects are

characterized through conceptual deepening (Márquez, 2014); and c) reflecting – thinking carefully to understand issues that allow forming an opinion and making decisions based on facts (Pérez and Merino, 2021).

Based on the documentary review, critical thinking is described as the liberating force in a civic person, fostering self-questioning and is capable of reflection, analysis, and investigation. In addition, this type of thinking stimulates curiosity, reason, flexibility, and impartiality to make judgments when analyzing environmental problems. In addition, logical thinking should be considered as part of the emergence of critical thinking, since Saiz and Rivas (2008) define the latter as a process of seeking knowledge through reasoning, problem-solving, and decision-making to achieve the desired results most efficiently. The above concept corresponds to that of Halpern (2014), who points out that critical thinking is more than just reflecting on one's thinking or making judgments and solving problems. Critical thinking uses evidence and reasoning, which is achieved by combining attitudes, knowledge, and thinking skills.

3.2 Validation of the PC-B test by a panel of experts

In order to correctly represent the items formulated in the PC-B test, the Content Validation Index (CVI) was employed, in which each test question is evaluated according to the criteria of essential, useful but not essential, and non-essential, as defined by Tristán-López (2008). It should be noted that the CVI is evaluated by experts who are familiar with the domain to be evaluated; therefore, the items of the questionnaire, or a bank of items, are submitted to the scrutiny of the experts, who classify them into the three categories mentioned above. The content validity ratio (CVR') is then determined as the proportion of agreements in the essential category. Then, items with a CVR' greater than or equal to 0.58 are considered acceptable. Finally, the unacceptable items are reviewed, corrected, and submitted again to the opinion of the panelists (Tristán-López, 2008).

The content validity was carried out by ten experts, who stand out for their level of training as doctors, experience in instrument evaluation and knowledge on the subject of critical thinking.

For the evaluation of the PC-B test questions, the following categories were taken into account: unsatisfactory (0): irrelevant level of critical thinking; satisfactory (1): relevant level of critical thinking; and superior (2): relevant level of critical thinking with comprehensive analysis and description. The test has thirty-six questions, divided into nine questions for each dimension.

Table 2. Operationalization of the construct

Construct	Dimension	Indicator	Items	Total of items
Critical thinking	Observation	Identifying	Preg_1 Preg_2 Preg_3	3
		Reporting	Preg_4 Preg_5 Preg_6	3
		Detecting	Preg_7 Preg_8 Preg_9	3
	Description	Analyzing	Preg_10	3

		Preg_11	
		Preg_12	
	Listing	Preg_13	
		Preg_14	3
		Preg_15	
	Reflecting	Preg_16	
		Preg_17	3
		Preg_18	
Comparison	Defining	Preg_19	
		Preg_20	3
		Preg_21	
	Relating	Preg_22	
		Preg_23	3
		Preg_24	
	Differentiating	Preg_25	
		Preg_26	3
		Preg_27	
Classification	Characterizing	Preg_28	
		Preg_29	3
		Preg_30	
	Proposing	Preg_31	
		Preg_32	3
		Preg_33	
	Associating	Preg_34	
		Preg_35	3
		Preg_36	

To establish the consensus of the experts, the Lawshe (1975) formula, cited in Tristán-López (2008), was used, which proposes the Content Validity Ratio (CVR). It is defined as:

$$CVR = \frac{n_e - \frac{N}{2}}{\frac{N}{2}}$$

Where:

ne = number of panelists (experts) who agree in the "essential" category.

N = total number of panelists (experts).

For the Content Validity Index (CVI), the following mathematical expression is used:

$$CVI = \frac{\sum_{i=1}^{M} CVRi}{M}$$

Where:

CVRi = Content Validity Ratio of acceptable items according to Lawshe's criteria.

M = Total number of acceptable test items.

Table 3 shows the values obtained in the PC-B test items in each category when applying the CVI. The items that were placed as non-essential were reviewed by the researchers. The evaluations were projected in the revision of the items from the contexts, wording, item coherence, and adjustments of the response options and suggestions in the location of dimensions and variables.

Table 3. Content validation index PC-B test.

Items	Essential	Useful but not essential	Non-essential	CVR Lawshe (1975)	CVR' Tristán-López (2008)	Is the item accepted?	CVI per Dimension	CVI
Preg_1	10	0	0	2,333	1,667	YES		
Preg_2	10	0	0	2,333	1,667	YES		
Preg_3	10	0	0	2,333	1,667	YES		
Preg_4	10	0	0	2,333	1,667	YES		
Preg_5	10	0	0	2,333	1,667	YES	1,59	
Preg_6	10	0	0	2,333	1,667	YES		
Preg_7	10	0	0	2,333	1,667	YES		
Preg_8	6	0	4	1,000	1,000	YES		
Preg_9	10	0	0	2,333	1,667	YES		
Preg_10	10	0	0	2,333	1,667	YES		
Preg_11	10	0	0	2,333	1,667	YES		
Preg_12	10	0	0	2,333	1,667	YES		
Preg_13	10	0	0	2,333	1,667	YES		1,5463
Preg_14	10	0	0	2,333	1,667	YES	1,59	
Preg_15	6	1	3	1,000	1,000	YES		
Preg_16	10	0	0	2,333	1,667	YES		
Preg_17	10	0	0	2,333	1,667	YES		
Preg_18	10	0	0	2,333	1,667	YES		
Preg_19	10	0	0	2,333	1,667	YES		
Preg_20	10	0	0	2,333	1,667	YES		
Preg_21	7	1	2	1,333	1,167	YES		
Preg_22	8	0	2	1,667	1,333	YES	1,50	
Preg_23	8	0	2	1,667	1,333	YES		
Preg_24	8	0	2	1,667	1,333	YES		
Preg_25	10	0	0	2,333	1,667	YES		

Preg_26	10	0	0	2,333	1,667	YES	
Preg_27	10	0	0	2,333	1,667	YES	
Preg_28	9	0	1	2,000	1,500	YES	
Preg_29	9	0	1	2,000	1,500	YES	
Preg_30	9	0	1	2,000	1,500	YES	
Preg_31	9	0	1	2,000	1,500	YES	
Preg_32	9	0	1	2,000	1,500	YES	1,50
Preg_33	9	0	1	2,000	1,500	YES	
Preg_34	10	0	0	2,333	1,667	YES	
Preg_35	7	1	2	1,333	1,167	YES	
Preg_36	10	0	0	2,333	1,667	YES	
Total amount of acceptable items					55,6666667	36	

CONTENT VALIDATION INDEX - PC-B TEST

Dimensions	Factor/indicator	Questions	Experts										Essential	Useful but not essential	Not essensial	Useful + Essential	RVC Thesis JV	RVC lawshe (1975)	RVC´ (Tristan, 2008)	Is the item accepted?	CVI por Dimensión	CVI
			E1	E2	E3	E4	E5	E6	E7	E8	E9	E10										
Observación	Identificar	1	1	1	1	1	1	1	1	1	1	1	10	0	0	10	1,0	2,3	1,7	SI		
		2	1	1	1	1	1	1	1	1	1	1	10	0	0	10	1,0	2,3	1,7	SI		
		3	1	1	1	1	1	1	1	1	1	1	10	0	0	10	1,0	2,3	1,7	SI		
	Dar a conocer	4	1	1	1	1	1	1	1	1	1	1	10	0	0	10	1,0	2,3	1,7	SI		
		5	1	1	1	1	1	1	1	1	1	1	10	0	0	10	1,0	2,3	1,7	SI	1,593	
		6	1	1	1	1	1	1	1	1	1	1	10	0	0	10	1,0	2,3	1,7	SI		
	Detectar	7	1	1	1	1	1	1	1	1	1	1	10	0	0	10	1,0	2,3	1,7	SI		
		8	1	1	1	-1	-1	-1	1	-1	1	1	6	0	4	6	0,2	1,0	1,0	SI		
		9	1	1	1	1	1	1	1	1	1	1	10	0	0	10	1,0	2,3	1,7	SI		
Comparación	Definir	10	1	1	1	1	1	1	1	1	1	1	10	0	0	10	1,0	2,3	1,7	SI		
		11	1	1	1	1	1	1	1	1	1	1	10	0	0	10	1,0	2,3	1,7	SI		
		12	1	1	1	1	1	1	1	1	1	1	10	0	0	10	1,0	2,3	1,7	SI		
	Relacionar	13	1	1	1	1	1	1	1	1	1	1	10	0	0	10	1,0	2,3	1,7	SI		
		14	1	1	1	1	1	1	1	1	1	1	10	0	0	10	1,0	2,3	1,7	SI	1,593	
		15	1	1	1	0	-1	-1	1	-1	1	1	6	1	3	7	0,4	1,0	1,0	SI		
	Diferenciar	16	1	1	1	1	1	1	1	1	1	1	10	0	0	10	1,0	2,3	1,7	SI		
		17	1	1	1	1	1	1	1	1	1	1	10	0	0	10	1,0	2,3	1,7	SI		
		18	1	1	1	1	1	1	1	1	1	1	10	0	0	10	1,0	2,3	1,7	SI		1,5462963
Clasificación	Caracterizar	19	1	1	1	1	1	1	1	1	1	1	10	0	0	10	1,0	2,3	1,7	SI		
		20	1	1	1	1	1	1	1	1	1	1	10	0	0	10	1,0	2,3	1,7	SI		
		21	1	1	1	0	-1	1	1	-1	1	1	7	1	2	8	0,6	1,3	1,2	SI		
	Proponer	22	1	1	1	-1	-1	1	1	1	1	1	8	0	2	8	0,6	1,7	1,3	SI		
		23	1	1	1	-1	-1	1	1	1	1	1	8	0	2	8	0,6	1,7	1,3	SI	1,50	
		24	1	1	1	-1	-1	1	1	1	1	1	8	0	2	8	0,6	1,7	1,3	SI		
	Asociar	25	1	1	1	1	1	1	1	1	1	1	10	0	0	10	1,0	2,3	1,7	SI		
		26	1	1	1	1	1	1	1	1	1	1	10	0	0	10	1,0	2,3	1,7	SI		
		27	1	1	1	1	1	1	1	1	1	1	10	0	0	10	1,0	2,3	1,7	SI		
Descripción	Analizar	28	1	1	1	-1	1	1	1	1	1	1	9	0	1	9	0,8	2,0	1,5	SI		
		29	1	1	1	-1	1	1	1	1	1	1	9	0	1	9	0,8	2,0	1,5	SI		
		30	1	1	1	-1	1	1	1	1	1	1	9	0	1	9	0,8	2,0	1,5	SI		
	Listar	31	1	1	1	-1	1	1	1	1	1	1	9	0	1	9	0,8	2,0	1,5	SI		
		32	1	1	1	-1	1	1	1	1	1	1	9	0	1	9	0,8	2,0	1,5	SI	1,5	
		33	1	1	1	-1	1	1	1	1	1	1	9	0	1	9	0,8	2,0	1,5	SI		
	Reflexionar	34	1	1	1	1	1	1	1	1	1	1	10	0	0	10	1,0	2,3	1,7	SI		
		35	1	1	1	-1	-1	1	1	0	1	1	7	1	2	8	0,6	1,3	1,2	SI		
		36	1	1	1	1	1	1	1	1	1	1	10	0	0	10	1,0	2,3	1,7	SI		
																			55,6667	36	1,5462963	

Figure 5. Item ratings by experts. CVI

After verifying that the items developed in the test obtained a value higher than 0.58, the average value of CVR', it is affirmed that the test has acceptable content validity, according to the experts who participated in the review.

Regarding the correlation between indicators, Spearman's correlation coefficient was used, also known as Spearman's rho, "which is used for quantitative variables with ordinal or freely distributed data" (Roy-García et al., 2019, pp. 355-356). The following formula was used for the calculation:

$$r_s = 1 - \frac{6\sum d^2}{n(n^2 - 1)}$$

Donde:

$\sum d^2$ = sum of the difference in ranks

N = number of pairs (X, Y).

The correlation coefficient can represent values between -1 and +1, where 0 means that there is no correlation, that is, that the variable is not related to the behavior of the other variable. A weak correlation corresponds to values more or less than 0.20, moderate when it reaches 0.50, good when it is located at 0.80, and perfect when the correlation is higher than 0.81. (Roy-García et al., 2019, p. 356).

3.3 PC-B validation test on the basis of the KMO test

To determine whether it is feasible to perform a factor analysis, the Kaiser-Meyer-Olkin (KMO) test was performed, which is a measure that determines how adequate the data are for such an analysis. It measures the adequacy of sampling for each indicator in the model. It also "indicates the proportion of variance in the variables that can be caused by underlying factors. High values (close to 1.0) generally indicate that a factor analysis may be useful with the data. If the value is less than 0.50, the results of the factor analysis will probably not be very useful" (International Business Machines [IBM], 2023). The frequency factor of the KMO test is given by the following equation:

$$\text{Call: KMO (r = mar\_cor); Overall MSA} = 0.7; \text{MSA per item}$$

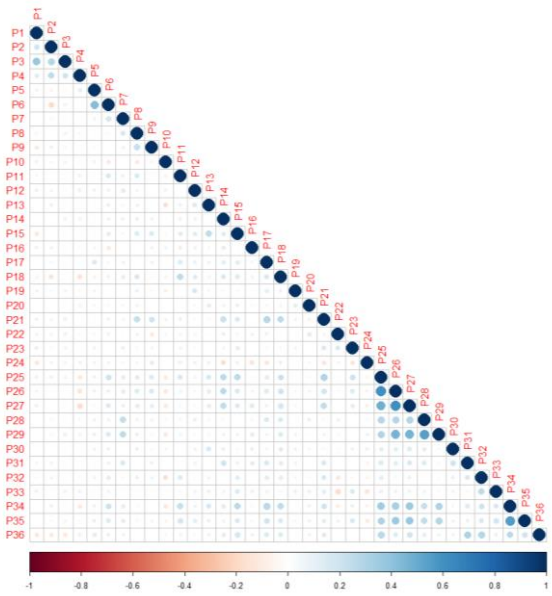

Figure 6. Correlation matrix

Figure 6 shows that the KMO frequency factor is 0.7, therefore, the factor analysis can be continued.

Table 4. Kaiser-Meyer-Olkin (KMO) Factor Adequacy

Preg_1	Preg_2	Preg_3	Preg_4	Preg_5	Preg_6	Preg_7	Preg_8	Preg_9	Preg_10
0.58	0.58	0.57	0.56	0.50	0.60	0.71	0.57	0.56	0.62
Preg_11	Preg_12	Preg_13	Preg_14	Preg_15	Preg_16	Preg_17	Preg_18	Preg_19	Preg_20
0.63	0.60	0.66	0.72	0.78	0.56	0.73	0.78	0.49	0.48

Preg_21	Preg_22	Preg_23	Preg_24	Preg_25	Preg_26	Preg_27	Preg_28	Preg_29	Preg_30
0.75	0.44	0.62	0.57	0.84	0.78	0.86	0.70	0.72	0.56
Preg_31	Preg_32	Preg_33	Preg_34	Preg_35	Preg_36				
0.67	0.72	0.62	0.80	0.79	0.78				

Additionally, Bartlett's test of sphericity was performed, which "tests the hypothesis that the correlation matrix is an identity matrix indicating that the variables are not related and, therefore, are not suitable for structure detection" (International Business Machines [IBM], 2023). This test specifies that if the significance level values are less than 0.05, then a factor analysis can be useful with the data. To establish the null hypothesis that states that the variables are not correlated in the population, the following equation is applied:

$$\text{cortest.bartlet}\left(\text{mat}_{\text{cor}}, \text{n} = \text{nrow(Data)}\right)\$\text{p [1]5.5944399e} - 96$$

The p-value result allows the null hypothesis to be rejected if it is less than 0.05, so factor analysis is suitable for use with the data. After checking that the data can be factored, Confirmatory Factor Analysis (FA) is applied, which is "a statistical data reduction technique used to explain correlations between observed variables in terms of a smaller number of unobserved variables called factors" (Parra, 2017).

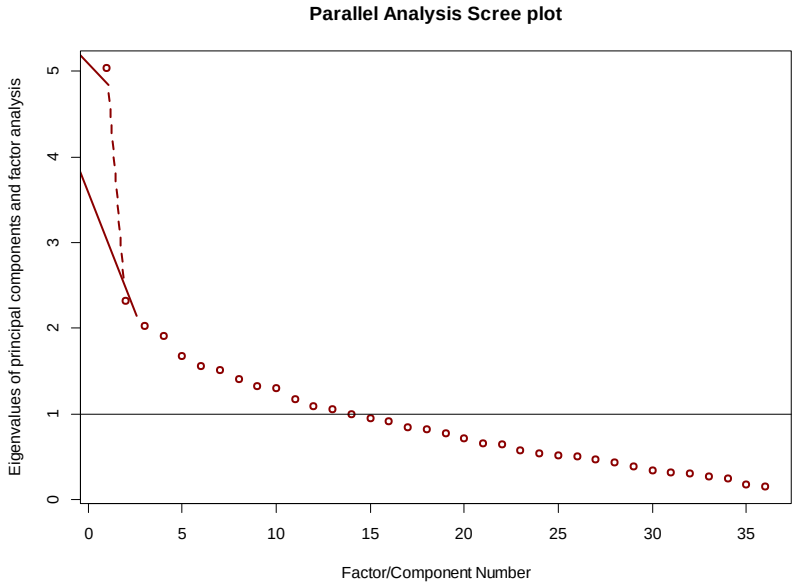

Figure 7. Factor selection by components

Table 5 shows the exploratory factor analysis (EFA) with the 12 factors.

Table 5. EFA with 12 factors

	PA1	PA7	PA3	PA2	PA5	PA4	PA8	PA6	PA10	PA12	PA9	PA11
P1	-0.135	-0.158	-0.209	0.295	0.225	0.333						
P2		-0.104		-0.275	0.108	0.456						
P3		-	-			0.470			0.291		-	

		0.160	0.122							0.151		
P4	-0.164		0.129	0.124	-0.175	0.697	0.129		-0.253	0.116		
P5		0.108		0.715							0.111	
P6	0.184			0.641	-0.108	-0.116				0.177	-0.137	0.228
P7				0.109								0.722
P8			0.113			0.101	0.718		0.250			
P9	0.241	-0.102		-0.144			0.373		-0.299			
P10	-0.158								-0.423			
P11		0.140	0.268						0.320	0.263		0.146
P12			0.284	0.193		0.134		-0.140	-0.177		0.359	0.173
P13	0.111		0.553								0.207	
P14	0.126		0.204		0.292		0.106			-0.113		
P15	0.149		0.681				0.177					
P16		0.135		-0.103						-0.445	0.204	
P17		0.324	0.150	0.193								-0.137
P18	0.120	0.182			0.118	-0.147	0.243		0.274	0.181	0.236	
P19		-0.113							0.150		0.420	-0.137
P20							-0.109	0.201	-0.100	0.176		
P21	0.117	0.193	0.101		0.293		0.465				0.117	-0.208
P22					0.121			0.885		-0.131		
P23	0.100				0.636			0.171		0.221		
P24		-0.105			-0.469	-0.117			-0.191		-0.126	
P25	0.660	0.103	0.243		0.257	-0.149	0.158			0.105		
P26	0.808		0.122		0.262		0.128	-0.205				
P27	0.785	0.130				-			0.140	-		

					0.146				0.175		
P28	0.433	0.226	-0.122		-0.205		-0.105	0.144	-0.209	0.196	0.195	0.180
P29	0.691				-0.275	0.161						0.128
P30	0.106	0.158									0.482	
P31		0.670	-0.113		0.126					-0.139		0.234
P32		0.411							0.260	0.149		-0.118
P33		0.148			0.110			-0.110		0.457	0.162	
P34	0.462	0.426	0.290			-0.105			0.151	0.250		
P35	0.529	0.235	0.232	0.150	0.137			0.119		0.183		
P36	0.162	0.256		0.116		-0.140						

	PA1	PA7	PA3	PA2	PA5	PA4	PA8	PA6	PA10	PA12	PA9	PA11
SS loadings	3.208	1.689	1.343	1.339	1.297	1.258	1.131	1.047	0.964	0.948	0.932	0.883
Proportion Var	0.089	0.047	0.037	0.037	0.036	0.035	0.031	0.029	0.027	0.026	0.026	0.025
Cumulative Var	0.089	0.136	0.173	0.211	0.247	0.281	0.313	0.342	0.369	0.395	0.421	0.446

Confirmatory factor analysis (CFA) "attempts to determine whether the number of factors obtained and their loadings correspond to those that would be expected in light of a prior theory about the data. The a priori hypothesis is that there are certain pre-established factors and that each of them is associated with a certain subset of the variables" (Parra, 2017).

Table 6. Statistical summary of the CFA

CHISQ	DF	SRMR	RMSEA	TLI	CFI
584.060	493.000	0.08	0.038	0.829	0.858

Note: CHISQ: Test statistic; DF: Degrees of freedom; SRMR: Standardized Root; Mean Square Residual; RMSEA: Root Mean Square Error of Approximation; TLI: Tucker-Lewis Index; CFI: Comparative Fit Index. Data obtained from the pilot test. Own elaboration.

6 specifies that the construct structure of the PC-B test is a relatively optimal model with a minimum Root Mean Square Error (RMSEA) and a high comparative fit index (CFI) and Tucker-Lewis Index (TLI) that are close to 1.

Figure 8 represents a descriptive map of the relationships of the questions associated with the dimensions and indicators presented in the PC-B test.

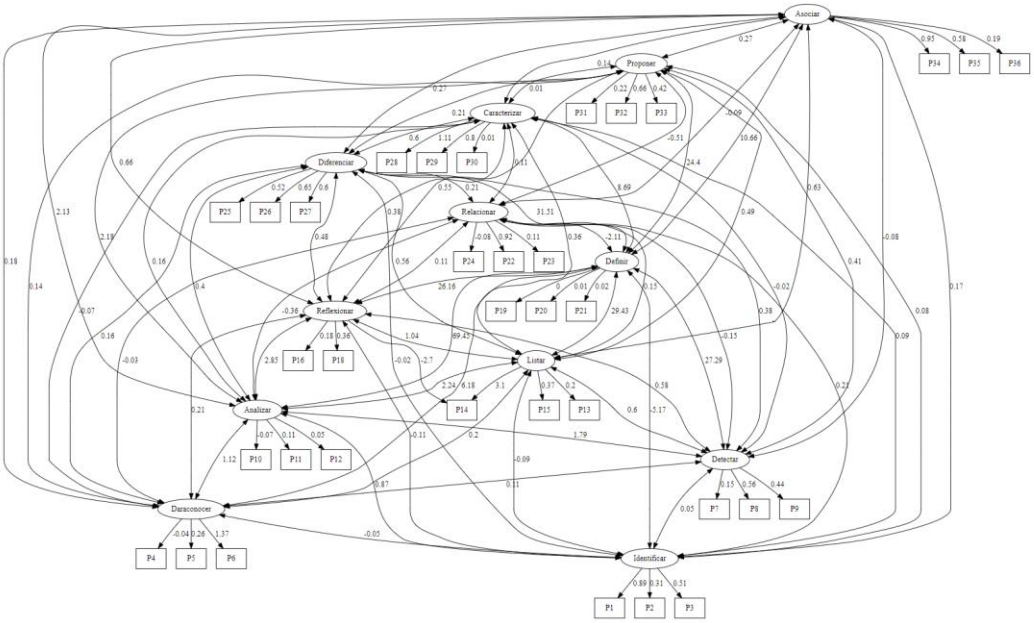

Figure 8. Graphical summary CFA

3.4 Degree of explanation of each item within the test structure

To establish the degree of explanation of the structure of the PC-B Test among the dimensions and indicators, the factorial scale or factorial saturation weight was used, which showed that the items with the highest degree of explanation in the structure of the PC-B Test are: P1, P2 and P34; while items P4, P6, P7, P10, P12, P13, P 19, P20, P21, P28 and P30 reached low degree of explanation. Table 7 shows the degrees of association of the PC-B test items.

Table 7. Degree of association of the items in the test structure

Item	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12
Estimate	0.79	0.095	0.257	0.002	0.069	NA	0.024	0.316	0.192	0.006	0.011	0.002
Item	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24
Estimate	0.041	-0.44	0.139	0.033	0	0.128	0	0	0	0.849	0.012	0.006
Item	P25	P26	P27	P28	P29	P30	P31	P32	P33	P34	P35	P36
Estimate	0.275	0.422	0.365	NA	0.646	0	0.051	0.434	0.181	0.902	0.336	0.034

Table 8 shows the factorial scales or factorial saturation weights that represent the degree to which each item explains the behavior of each question associated with the indicator to determine the relationship of the item to its indicator.

Table 8. Representative factorial scale of items and variables

Latent Variables	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
Identifying						
P1	0.886	0.139	6.388	0.000	0.886	0.889
P2	0.307	0.100	3.072	0.002	0.307	0.308
P3	0.505	0.108	4.677	0.000	0.505	0.507
Reporting						

P4	-0.040	0.065	-0.619	0.536	-0.040	-0.041
P5	0.262	0.150	1.739	0.082	0.262	0.263
P6	1.368	0.647	2.115	0.034	1.368	1.373
Detecting						
P7	0.154	0.110	1.390	0.164	0.154	0.154
P8	0.561	0.124	4.536	0.000	0.561	0.562
P9	0.437	0.112	3.902	0.000	0.437	0.438
Analyzing						
P10	-0.075	0.336	-0.222	0.824	-0.075	-0.075
P11	0.105	0.467	0.225	0.822	0.105	0.105
P12	0.046	0.209	0.220	0.826	0.046	0.046
Listing						
P13	0.203	0.099	2.038	0.042	0.203	0.203
P14	3.094	5.570	0.555	0.579	3.094	3.098
P15	0.372	0.111	3.356	0.001	0.372	0.373
Reflecting						
P16	0.180	0.096	1.868	0.062	0.180	0.181
P17	-2.693	5.581	-0.483	0.629	-2.693	-2.697
P18	0.356	0.123	2.889	0.004	0.356	0.358
Defining						
P19	0.002	1.081	0.002	0.098	0.002	0.002
P20	0.006	2.626	0.002	0.098	0.006	0.006
P21	0.015	7.143	0.002	0.098	0.015	0.015
Relating						
P22	0.918	0.507	1.810	0.070	0.918	0.921
P23	0.111	0.108	1.027	0.305	0.111	0.112
P24	-0.075	0.101	-0.740	0.459	-0.075	-0.075
Differentiating						
P25	0.523	0.089	5.881	0.000	0.523	0.524
P26	0.649	0.087	7.470	0.000	0.649	0.650
P27	0.603	0.088	6.890	0.000	0.603	0.604
Characterizing						
P28	1.107	0.068	16.239	0.000	1.107	1.111
P29	0.801	0.078	10.254	0.000	0.801	0.804
P30	0.011	0.072	0.151	0.880	0.011	0.011
Proposing						

P31	0.224	0.106	2.115	0.034	0.224	0.225
P32	0.657	0.119	5.542	0.000	0.657	0.659
P33	0.424	0.104	4.062	0.000	0.424	0.425
Associating						
P34	0.947	0.109	8.710	0.000	0.947	0.950
P35	0.577	0.096	6.001	0.000	0.577	0.579
P36	0.185	0.092	2.003	0.045	0.185	0.185

According to the results obtained, it is established that the indicators that make up the PC-B test are statistically adequately representative. The following is the detail of the degree of representativeness between indicator items.

- Identifying: all 3 items are relatively adequately represented.
- Reporting: item P6 is relatively adequately represented, while items P4 and P5 present a low degree of explanation.
- Detecting: items P8 and P9 are relatively adequately represented, while item P7 presents a low degree of explanation.
- Analyzing: the three items present a low degree of explanation; it is inferred that the students present difficulty in solving these questions.
- Listing: items P14 and P15 are relatively well represented, while item P13 presents a low degree of explanation.
- Reflecting: items P17 and P18 are relatively well represented, while item P16 shows a low degree of explanation.
- Define: all three items present a low degree of explanation.
- Relating: item P22 is relatively well represented, while items P23 and P24 present a low degree of explanation.
- Differentiate: all three items present a good degree of explanation.
- Characterizing: items P28 and P29 are relatively well represented, while item P30 shows a low degree of explanation.
- Proposing: items P32 and P33 are relatively adequately represented, while item P31 has a low degree of explanation.
- Associating: items P34 and P35 are relatively adequately represented, while item P36 shows a low degree of explanation.

From the analysis of the items and indicators, there are relationships in the explanation that students make, according to the items, in an adequate response based on the statement and a low explanation in the descriptors presented.

Degree of explanation and relationship between dimensions and indicators

Table 9 shows the covariances between the dimensions and the latent indicators of the PC-B test. It can be seen that the latent variables that show the highest correlation between them are identify and characterize; make known and detect; characterize and define; detect and characterize; analyze and characterize; list and characterize; reflect and characterize; define and characterize; relate and characterize; differentiate and characterize; characterize and associate; propose and associate.

Table 9. Covariances between dimensions or latent variables

Covariances	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
Identify						
Reporting	0.272	0.041	6.548	0.000	0.164	0.164
Detecting	-0.027	0.029	-0.916	0.360	-0.074	-0.074
Analyzing	-0.009	0.024	-0.368	0.713	-0.064	-0.064
Listing	-0.025	0.014	-1.765	0.078	-0.163	-0.163
Reflecting	-0.032	0.019	-1.677	0.094	-0.247	-0.247
Defining	-0.006	0.016	-0.392	0.695	-0.076	-0.076
Relating	0.010	0.017	0.577	0.564	0.381	0.381
Differentiating	-0.008	0.015	-0.557	0.578	-0.030	-0.030
Characterizing	0.572	0.061	9.351	0.000	1.112	1.112
Proposing	0.043	0.023	1.874	0.061	0.173	0.173
Associating	0.002	0.023	0.078	0.938	0.005	0.005
Reporting						
Detecting	0.712	0.081	8.775	0.000	0.223	0.223
Analyzing	-0.364	0.071	-5.168	0.000	-0.293	-0.293
Listing	-0.202	0.041	-4.962	0.000	-0.149	-0.149
Reflecting	-0.231	0.052	-4.436	0.000	-0.206	-0.206
Defining	-0.436	0.051	-8.580	0.000	-0.606	-0.606
Relating	0.065	0.048	1.366	0.172	0.280	0.280
Differentiating	0.493	0.048	10.283	0.000	0.201	0.201
Characterizing	0.947	0.073	12.986	0.000	0.210	0.210
Proposing	0.225	0.062	3.611	0.000	0.103	0.103
Associating	-0.190	0.064	-2.976	0.003	-0.060	-0.060
Detecting						
Analyzing	-0.088	0.061	-1.448	0.148	-0.322	-0.322
Listing	0.107	0.035	3.012	0.003	0.360	0.360
Reflecting	0.089	0.046	1.934	0.053	0.363	0.363
Defining	0.146	0.040	3.628	0.000	0.928	0.928
Relating	0.156	0.044	3.509	0.000	3.059	3.059
Differentiating	0.195	0.039	5.014	0.000	0.363	0.363
Characterizing	1.284	0.072	17.875	0.000	1.297	1.297
Proposing	-0.024	0.055	-0.439	0.661	-0.051	-0.051
Associating	0.335	0.059	5.679	0.000	0.484	0.484
Analyzing						

Listing	-0.101	0.030	-3.339	0.001	-0.869	-0.869
Reflecting	-0.113	0.039	-2.864	0.004	-1.173	-1.173
Defining	-0.170	0.036	-4.736	0.000	-2.770	-2.770
Relating	-0.044	0.036	-1.230	0.219	-2.233	-2.233
Differentiating	-0.071	0.032	-2.220	0.026	-0.336	-0.336
Characterizing	-0.437	0.061	-7.116	0.000	-1.131	-1.131
Proposing	-0.182	0.049	-3.742	0.000	-0.976	-0.976
Associating	-0.239	0.051	-4.684	0.000	-0.882	-0.882
Listing						
Reflecting	0.128	0.035	3.673	0.000	1.224	1.224
Defining	0.111	0.026	4.275	0.000	1.653	1.653
Relating	-0.025	0.021	-1.211	0.226	-1.146	-1.146
Differentiating	0.196	0.031	6.425	0.000	0.857	0.857
Characterizing	0.688	0.064	10.795	0.000	1.634	1.634
Proposing	0.162	0.030	5.355	0.000	0.792	0.792
Associating	0.220	0.032	6.765	0.000	0.745	0.745
Reflecting						
Defining	0.033	0.026	1.296	0.195	0.603	0.603
Relating	-0.024	0.027	-0.885	0.376	-1.343	-1.343
Differentiating	0.106	0.027	3.991	0.000	0.560	0.560
Characterizing	0.867	0.067	13.026	0.000	2.486	2.486
Proposing	0.165	0.037	4.394	0.000	0.976	0.976
Associating	0.242	0.040	6.064	0.000	0.989	0.989
Defining						
Relating	0.041	0.024	1.736	0.083	3.558	3.558
Differentiating	0.320	0.032	9.897	0.000	2.636	2.636
Characterizing	0.614	0.052	11.745	0.000	2.750	2.750
Proposing	0.189	0.034	5.579	0.000	1.747	1.747
Associating	0.180	0.034	5.370	0.000	1.153	1.153
Relating						
Differentiating	0.209	0.028	7.396	0.000	5.327	5.327
Characterizing	0.788	0.066	11.883	0.000	10.909	10.909
Proposing	-0.153	0.036	-4.311	0.000	-4.387	-4.387
Associating	-0.018	0.034	-0.533	0.594	-0.360	-0.360
Differentiating						

Characterizing	0.570	0.044	12.824	0.000	0.747	0.747
Proposing	0.357	0.039	9.097	0.000	0.966	0.966
Associating	0.368	0.037	9.898	0.000	0.688	0.688
Characterizing						
Proposing	-0.108	0.050	-2.140	0.032	-0.158	-0.158
Associating	0.983	0.064	15.469	0.000	1.000	1.000
Proposing						
Associating	0.274	0.048	5.761	0.000	0.577	0.577

3.5 Instrument

This research describes the contextualized Critical Thinking Test called the PC-B Test, built by the authors of this research, which aims to identify the levels of critical thinking of first-year students. Four dimensions are presented with their indicators extracted from the theoretical review of several authors who have proposed different tests to evaluate critical thinking in students, the social context was considered to assess the subjects around the reasoning that is determined in everyday scenarios and knowledge.

The PC-B Test presents a categorization of responses according to ordinal variables stratified as superior valued at two (2), Satisfactory valued at one (1), and Unsatisfactory valued at zero (0). The first category of valuation establishes if the individual reaches higher levels of critical thinking, the second category locates those individuals who possess some of the dimensions associated with the PC-B Test, while the last category points to individuals who lack elemental capacities from the Observation, Comparison, Classification, and Description.

The PC-B test consists of four configurations of questions, ranging from multiple choice, associations, relations, and deductive relations, where the student projects his answer according to the contexts presented. The test can be applied individually or collectively in a maximum time of two hours. The following graphics show the structure of the questions.

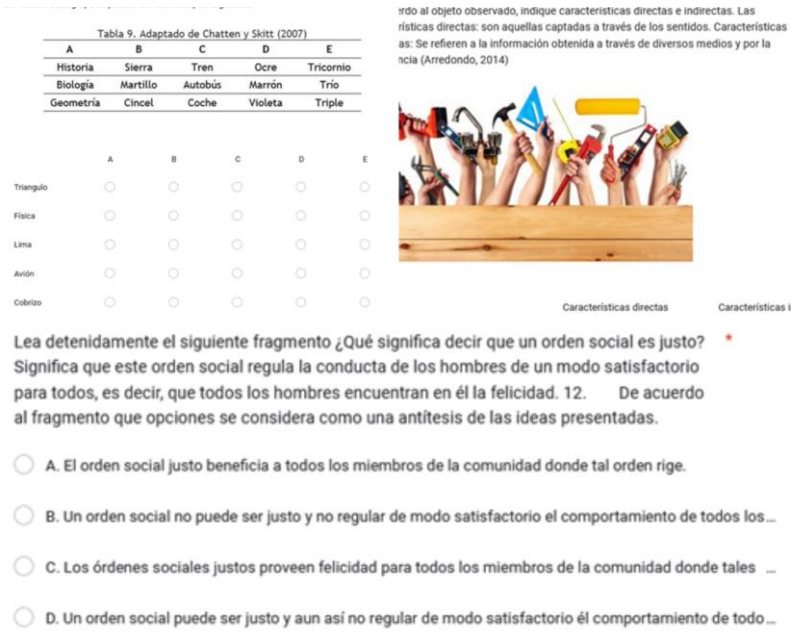

Figure 9. Categorization of contextualized questions of PC-B test

Regarding the interpretation of the PC-B test, the quantitative scores obtained by the students are considered to establish a diagnosis from the PC levels in the dimensions covered by the test. The highest score a student can obtain is seventy-two points (72), distributed as shown in table 10.

Table 10. Distribution of ranges determining PC assessment in the basic dimensions established in the PC-B Test

PC level	Point ranges
Superior	When individuals obtain scores that fall within the range of 49 to 72 points.
Satisfactory	When the individual obtains ratings that are in the 25-to-48-point range
Unsatisfactory	When the individual obtains ratings that are between 1 to 24 points

4. Discussions and conclusions

The purpose of the PC-B Test is to evaluate critical thinking at a basic level in first-year university students based on the dimensions and indicators that compose it. This evaluation aims to pinpoint strategies that bolster advanced cognitive abilities in students, leading to enhanced competencies in knowledge and thinking. In addition, the results of the application of the test will serve to articulate the curricular plan of the different university programs in terms of the improvement of critical thinking, through the formulation and implementation of competencies and learning outcomes that will lead students to move up the development scale of critical thinking until they graduate from the university. From the national perspective, in university and secondary education, processes are oriented towards achieving critical thinking in students to form a citizen with the capacity to solve problems of the context.

The validation of the PC-B test was carried out by a panel of experts who evaluated the test items according to the qualitative categories, based on the CVI technique proposed by Tristán-López (2008). The results of the validation determined the categorical organization of the dimensions and associated indicators from a basic order to a higher order. Similarly, questions were formulated based on the less privileged contexts; socioeconomic, cultural, and socio-demographic conditions of the students.

It was subjected to the Kaiser-Meyer-Olkin (KMO) test, indicating the possibility of factor analysis validation. In addition, Barlett's test of sphericity was used for hypothesis testing, which suggested that the PC-B test could be subjected to exploratory factor analysis (EFA) to determine the correlations between items, indicators, and dimensions and confirm the relationships between variables and indicators made the results reliable.

By establishing the degree of explanation of the test structure through confirmatory factor analysis (CFA), it was determined the factorial saturation weights between indicators and items, and the questions with the highest weight according to the indicators were identified. To measure the degree of relationship between variables and covariances, strong relationships were found between indicators such as: identifying and characterizing; reporting and detecting; characterizing and defining; detecting and characterizing; analyzing and characterizing; listing and characterizing; reflecting and characterizing; defining and characterizing; relating and characterizing; differentiating and characterizing; characterizing and associating; preparing and associating. It demonstrates that the PC-B test has a cohesive and coherent structure according to the established categories and contextualized questions.

In the pilot test applied to a group of first-year students from universities in the Colombian Caribbean region, it was evident that the items proposed in the Test-PCB had

cohesion and coherence between the dimensions and indicators formulated since the students answered the contextualized questions clearly and unambiguously.

References

Alquichire R., S. L., & Arrieta R., J. C. (2018). Relación entre habilidades de pensamiento crítico y rendimiento académico. Voces Y Silencios. Revista Latinoamericana De Educación, 9(1), 28–52. https://doi.org/10.18175/vys9.1.2018.03

Arredondo, M. (2014). Habilidades básicas para aprender a pensar. Trillas.

Bigg, J. (2006). Calidad del aprendizaje Universitario. Narcea.

Calle, Y. (2014). Las habilidades del pensamiento crítico durante la escritura digital en un ambiente de aprendizaje apoyado por herramientas de la web 2.0. Encuentros, 12(1), 27-45. http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S1692-58582014000100003&lng=en&tlng=es.

Carrasco, J. (2018). Medición del desarrollo del pensamiento crítico en estudiantes chilenos/as de educación superior. [Tesis de maestría, Universidad de Concepción]. Repositorio institucional - Universidad de Concepción. http://repositorio.udec.cl/handle/11594/3062

Ennis, R., & Weir, E. (1985). The Ennis-Weir Critical Thinking Essay Test. Midwest Publications. http://evolkov.net/critic.think/tests/Ennis-Weir.Critic.Think.Essay.Test.pdf

Ennis, R., & Millman, J. (2005a). Cornell Critical Thinking Test Level X. Critical Thinking Company. https://www.criticalthinking.com/cornell-critical-thinking-test-level-x.html

Ennis, R.H., Millman, J., & Tomko, T. N. (1985). Cornell Critical Thinking Tests, Level X and Level Z Manual (3rd ed.). Pacific Grove, Midwest Publications.

Ennis, R. H., Millman, J., & Tomko, T.N. (2005). Cornell Critical Thinking Tests. Seaside, The Critical Thinking Co.

Ennis, R., Millman, J., & Tomko, T. (2012). Cornell Critical Thinking Test Level Z. Critical Thinking Company. https://www.criticalthinking.com/cornell-critical-thinking-test-level-z.html

Escobedo, M., Hernández, J., Ortega, E. y Martínez, G. (2016). Modelos de ecuaciones estructurales: Características, fases, construcción, aplicación y resultados. Ciencia & trabajo, 18(55), 16-22. https://dx.doi.org/10.4067/S0718-24492016000100004

Facione, P. (1991). Using the California Critical Thinking Skills Test in research, evaluation, and assessment. ERIC. https://eric.ed.gov/?id=ED337498

Facione, N., Facione, P., & Sánchez, C. (1994). Critical thinking disposition as a measure of competent clinical judgment: The development of the California Critical Thinking Disposition Inventory. The Journal of Nursing Education, 33(8), 345–350. https://pubmed.ncbi.nlm.nih.gov/7799093/

García, M., Acosta, D., Atencia, A., y Rodríguez, 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). https://doi.org/10.6018/reifop.435831

Soca, R. (2007). Detectar. En Etimología: El Origen de las Palabras. https://www.ellibrototal.com/ltotal/

Gyenes, A. (2021). Student perceptions of critical thinking in EMI programs at Japanese universities: A Q-methodology study. Journal of English for Academic Purposes, 54. https://doi.org/10.1016/j.jeap.2021.101053

Halpern, D. (2006). Halpern Critical Thinking Assessment Using Everyday Situations: Background and scoring standards [Unpublished Manuscript]. Claremont McKenna College.

Halpern, D. (2014). Thought and knowledge: An introduction to critical thinking. Psychology Press.

Hernández-Sampieri, R. y Mendoza, Ch. (2018). Metodología de la investigación: las rutas cuantitativa, cualitativa y mixta. McGraw-Hill.

Hyytinen, H., Silvennoinen, K., Kleemola, K., & Toom, A. (2021). The dynamic relationship between response processes and self-regulation in critical thinking assessments. Studies in Educational Evaluation, 71, 1-12. https://doi.org/10.1016/j.stueduc.2021.101090

International Business Machines IBM. (2023). Prueba de KMO y Bartlett. https://www.ibm.com/docs/es/spss-statistics/29.0.0?topic=detection-kmo-bartletts-test

Kamin, C. S., O'Sullivan, P. S., Younger, M., & Deterding, R. (2001). Measuring critical thinking in problem-based learning discourse. Teaching and Learning in Medicine, 13(1), 27-35. https://doi.org/10.1207/S15328015TLM1301_6

Márquez, P. (2014). Cómo desarrollar habilidades de pensamiento. Guía basada en cinco propuestas de investigadores. Ediciones de la U.

Marzano, R. (2001). Designing a new taxonomy of educational objectives. Experts in Assessment. Corwin Press Inc. https://eric.ed.gov/?id=ED447161

Norris, S. y Ennis, R. (1989). Evaluating Critical Thinking. The Practitioners' Guide to Teaching Thinking Series. Pacific Grove: Midwest Publications. https://eric.ed.gov/?id=ed404836

Ossa-Cornejo, C., Palma-Luengo, M., Lagos-San Martín, N., Quintana-Abello, I., y Díaz-Larenas, C. (2017). Analysis of critical thinking measuring instruments. Ciencias Psicológicas, 11(1), 19-28. http://dx.doi.org/10.22235/cp.v11i2.1343

Paul, R., & Elder, L. (2019). The International Critical Thinking Reading and Writing Test. Rowman & Littlefield Publishers. https://www.criticalthinking.org/store/products/the-international-critical-thinking-readingwriting-test/229

Scriven, M. & Paul, R. (2003). Defining critical thinking. https://www.criticalthinking.org/template.php?pages_id=766

Parra, F. (2017). Estadística y machine learning con R. https://rpubs.com/PacoParra/293407

Pérez Porto, J., Gardey, A. (10 de junio de 2023). Caracterizar - Qué es, definición y concepto. Definicion.de. Última actualización el 23 de octubre de 2017. Recuperado el 10 de junio de 2023 de https://definicion.de/caracterizar/

Pérez, J., Merino, M. (10 de junio de 2023). Propuesta - Qué es, definición y concepto. Definicion.de. Última actualización el 18 de enero de 2023. Recuperado el 10 de junio de 2023 de https://definicion.de/propuesta/

Pérez, J., Merino, M. (10 de junio de 2023). Reflexión - Qué es, definición, en la informática y en la geometría. Definicion.de. Última actualización el 11 de octubre de 2021. Recuperado el 10 de junio de 2023 de https://definicion.de/reflexion/

Polo, B., Ramírez, G., Hinojosa, C. y Castañeda, W. (2022). Competencias transversales en el contexto educativo universitario: un pensamiento crítico desde los principios de la gamificación, Revista Prisma Social, (38), pp. 159-178. https://revistaprismasocial.es/issue/view/241

Rivas, S. y Saiz, C. (2012). Validación y propiedades psicométricas de la prueba de pensamiento crítico PENCRISAL. Rema, 17(1), 18-34. https://dialnet.unirioja.es/servlet/articulo?codigo=4107460

Roca, J. (2013). El desarrollo del pensamiento crítico a través de diferentes metodologías docentes en el Grado en Enfermería. [Tesis doctoral, Universitat Autónoma de Barcelona]. https://dialnet.unirioja.es/servlet/tesis?codigo=84936

Roy-García, I., Rivas-Ruiz, R., Pérez-Rodríguez, M. y Palacios-Cruz, L. (2019). Correlación: no toda correlación implica causalidad. Revista Alergia México, 66(3), 354-360. https://doi.org/10.29262/ram.v66i3.651

Saiz, C., y Rivas, S. (2008). Evaluación del pensamiento crítico: una propuesta para diferenciar formas de pensar, Ergo, Nueva Epoca, (22). http://cdigital.uv.mx/handle/123456789/38320

Saiz, C., y Rivas, S. (2011). Evaluation of the ARDESOS program: An initiative to improve critical thinking skills, Journal of the Scholarship of Teaching and Learning, 11(2), pp. 34 - 51. https://scholarworks.iu.edu/journals/index.php/josotl/issue/view/167

Sánchez, A. (2010). Introducción: ¿qué es caracterizar? Fundación Universitaria Católica del Norte. https://www.ucn.edu.co/sistema-investigacion/Documents/instrumento%20para%20caracterizar%20experiencias.pdf

Sánchez, H. (2013). La comprensión lectora, base del desarrollo del pensamiento crítico: Segunda parte. Horizonte de la Ciencia, 3(5), 31-38. https://dialnet.unirioja.es/servlet/articulo?codigo=5420526

De Sánchez, M. A. (1995). Desarrollo de habilidades de pensamiento: procesos básicos del pensamiento. Trillas, ITESM.

Strauss A. y Corbin J. (2002). Bases de la investigación cualitativa. Técnicas y procedimientos para desarrollar la teoría fundamentada. Editorial Universidad de Antioquia

Tristán-López, A. (2008). Modificación al modelo de Lawshe para el dictamen cuantitativo de la validez de contenido de un instrumento objetivo. Avances en medición, 6(1), 37-48. https://www.humanas.unal.edu.co/lab_psicometria/application/files/9716/0463/3548/VOL_6._Articulo4_Indice_de_validez_de_contenido_37-48.pdf

Valenzuela, J. y Nieto, A. (2008). Motivación y pensamiento crítico: aportes para el estudio de esta relación. Revista Electrónica de Motivación y Emoción, 11(28). http://reme.uji.es/articulos/numero28/article3/article3.pdf

Valenzuela, J., Nieto, A. y Saiz, C. (2011). Critical Thinking Motivational Scale (CTMS): una aportación para el estudio de la relación entre el pensamiento crítico y la motivación. Electronic Journal of Research in Educational Psychology, 9(2), 823-848. https://doi.org/10.25115/ejrep.v9i24.1475

Velásquez, B., Remolina, N. y Calle, M. (2013). Habilidades de pensamiento como estrategia de aprendizaje para los estudiantes universitarios. Revista De Investigaciones UNAD, 12(2), 23–41. https://doi.org/10.22490/25391887.1174

Watson, G. y Glaser, E. (1981). Watson-Glaser Critical Thinking Appraisal. Psychological Corporation. https://www.assessmentday.co.uk/watson-glaser-critical-thinking.htm

Watson, G. y Glaser, E. (2002). Watson – Glaser Critical Thinking Appraisal – UK Edition. Pearson.