JSSER
www.jsser.org

Journal of Social Studies Education Research

Sosyal Bilgiler Eğitimi Araştırmaları Dergisi

2024:15 (3), 213-228

eComplexity: Psychometric properties to test the validity and reliability of the instrument

Isolda Margarita Castillo-Martínez[1], Davis Velarde-Camaqui[2], María Soledad Ramírez-Montoya[3] & Jorge Sanabria-Z[4]

Abstract

Reasoning for complexity is a fundamental competency in these complex times for solutions to social problems and decision-making. The purpose of this paper is to demonstrate the validity and reliability of the eComplexity instrument by presenting its psychometric properties. The instrument consists of a Likert-type scale questionnaire designed to measure college students' perceptions of their mastery levels of complex reasoning competency as well as the subcompetencies that comprise it. The instrument was applied to higher education students in different countries of Latin America and Spain. The questionnaire articulates four types of thinking: systemic, scientific, critical and innovative and consists of 25 items. The methodology used was instrumental and psychometric, which seeks to demonstrate the validity and reliability of the eComplexity questionnaire. The results obtained from the research were as follows: The exploratory factor analysis indicated a Kaiser-Meyer-Olkin index (KMO) > .80, a significance of p< .05 and a Cronbach's Alpha value of 0.93. Likewise, a Confirmatory Factor Analysis was carried out and was possible to corroborate the internal structure validity of the instrument. In addition to Cronbach's Alpha coefficient, McDonald's Omega, and Guttman's Lambda coefficients were calculated to calculate reliability. With the results obtained it was possible to conclude that the instrument is valid and reliable, can be used in various university contexts to support integrated training necessary to address current challenges and contribute to educational research. It is recommended for future studies that the research can be expanded by using an instrument that can move from perceptual terms to measuring levels of complex reasoning mastery, but it is valuable to contrast with the students' perception to have a broader vision.

Keywords: *Complex reasoning, educational innovation, higher education. Likert scale, psychometric properties*

Introduction

The complexity paradigm is opposed to the simplification paradigm. The latter observes the object of knowledge disjunctively, identifying linearity in a cause-effect relationship. On the contrary, the complexity paradigm proposes to use a multifaceted and multidisciplinary approach (IGI

[1] *Dr., School of Humanities and Education, Tecnologico de Monterrey, Mexico; Email: isoldamcm@hotmail.com*
[2] *Dr., Institute for the Future of Education, Tecnologico de Monterrey, Mexico; Email: a00832505@tec.mx*
[3] *Dr., Institute for the Future of Education, EGADE Business School, Tecnologico de Monterrey, Mexico; Email: solramirez@tec.mx*
[4] *Dr., Department of Design, School of Architecture, Art and Design (EAAD), Tecnologico de Monterrey; Email: jorge.sanabria@tec.mx*

Global, 2021), analyzing phenomena as an indivisible whole. It contemplates seven principles regarding the object of knowledge: it is systemic, holographic, retroactive, recursive (what is cause is effect, but also vice versa), dialogic (everything happens concurrently although the parts may be contradictory), and the observation of the knowing subject is reintroduced (Morin, 2015). Complex thinking implies a strategy inseparable from the inventive participation of those who develop it. It is necessary to methodologically test (by walking) the generating principles of the method and simultaneously invent and create new principles (Morin, 2006). In this era, students must avoid a fragmented vision; they must develop complex reasoning to face current challenges. It is essential to identify what is involved in reasoning for complexity. The competence of reasoning for complexity allows the student to apply integrated thinking that facilitates analysis, synthesis, problem solving and continuous learning to master the cognitive skills necessary for scientific, critical and systemic thinking (Tecnologico de Monterrey, 2019). In addition to the three types of thinking considered for complex reasoning in the Guiding Document for Higher Education Teachers (Tecnologico de Monterrey, 2019), the eComplexity instrument analyzed in this article integrates another type of thinking, innovative thinking, since it allows a more comprehensive thinking, corresponding to complex reasoning.

This competence comprises four subcompetencies: a) Systemic thinking allows solving problems by interpreting data from different fields of science and determining the relevance of the elements of the system by analyzing the existing data set (Izvorska, 2016; Jaaron and Backhouse, 2018); b) Scientific thinking allows solving problems and questioning reality, objectively with valid and reliable methodologies, analyzing data to qualify the veracity. It also encompasses cognitive processes, such as inductive and deductive reasoning, problem solving, and hypothesis formulation and testing (Koerber et al., 2015; Suryansyah et al., 2021; Zimmerman and Croker, 2014); c) Critical thinking is the intellectually disciplined process of actively and skillfully conceptualizing, applying, analyzing, synthesizing, and evaluating information obtained or generated by observation, experience, reflection, reasoning, or communication to guide belief and action (Sellars et al., 2018; Straková and Cimermanová, 2018); and finally d) Innovative thinking incorporates some or all the following traits such as curiosity, imagination, inventiveness, playfulness, flexibility, and persistence. Integrated thinking enables the learner to meet challenges academically and in any domain of his or her life.

To provide solutions to the problems of our time, we must have a way of thinking in accordance with today's complex and changing environment. The reform of thought is a necessary condition to be subjects in humanity. An autonomous subject is only possible if there is a reform of thought (Morin, 2003). The author mentions that the reform of thought is necessary to articulate and organize knowledge and thus recognize and know the problems of the world. For knowledge to be relevant, education must provide evidence: a) Context: Information and elements must be contextually located to acquire meaning. b) The global: A society is more than a context; it is an organizing whole of which we are part. c) The multidimensional: Multidimensional information must comprise relevant knowledge. d) The complex: Unity and multiplicity must be combined (Morin, 2015). Practicing complexity requires continuously seeking knowledge and self-knowledge with no deadline for the cessation of effects. It involves conceiving to what extent thoughts, ghosts and ideas dominate, or whether one has the freedom to dream, change one's thoughts, break with ghosts and change one's ideas (Morin, 2021). Developing complex thinking or reasoning involves reforming our thinking to adapt it to the needs and problems that arise in our environment.

This article aims to answer the following question: What is the evidence of validity and reliability of the eComplexity instrument? We considered it essential to conduct a literature review to identify articles related to complex reasoning competence that described research in which questionnaires were used to measure it. We analyzed the database supporting the article "Title" (Author1). Of the 35 articles reviewed, we were only able to identify five studies that used questionnaires as instruments in the research process (link). In the study entitled "Defining and measuring innovative thinking among engineering undergraduates" (Amelink et al., 2013), the Modified Learning Strategies Questionnaire (MSLQ) was used for the part related to learning behaviors, which provided a measure of skills linked to innovative thinking. According to the authors of the research, the scale presented a moderate to acceptable degree of reliability (α .70) by Cronbach's alpha scores (Furr & Bacharach, 2014; Lord & Novick, 2008). In the article "Critical thinking skills of engineering students: undergraduate vs. graduate students" (Douglas, 2006), the California Critical Thinking Skills Test (CCTST) Form 2000 was used to measure critical thinking skills. It consisted of 34 multiple-choice items. The reliability of the Form 2000 was approximately 0.8. In the article "How and when is complex reasoning produced? Empirically driven development of a learning progression focused on complex reasoning about biodiversity" (Songer et al., 2009), the information functions of their test curves had IRT reliabilities of 0.78, 0.67 and 0.64 for all items, complex items and standardized items, respectively. The results showed that all items sufficiently fit their model. The mapped categories of inquiry items (minimal, intermediate, and complex) generally corresponded to the difficulty levels of the minimal, intermediate, and complex items. In the research entitled "Course-Embedded Undergraduate Research Experiences: The Power of Strategic Course Design" (Powell y Harmon, 2016), student perceptions were measured by administering a modified ROLE survey at the end of each course. Modifications included leaving spaces for open-ended reflective comments and the addition of four Likert-scale items addressing student outcomes. The article does not address the reliability aspect of the instrument used. Nor was test-retest reliability identified for the instruments used in the research "Understanding Differences in Student Learning Outcomes between the U.S. and Norwegian Educational Systems" (Safari et al., 2019), in which the instruments administered were the Herrmann Brain Dominance Inventory (HBDI), the Longitudinal Assessment of Self-Efficacy in Engineering (LAESE) and the Thinking Assessment Test.

The contribution of the present research is to complement the design and validation process of the eComplexity instrument -which measures the perception of university students regarding their levels of mastery of the complex reasoning competence, considering four sub competencies: systemic thinking, scientific thinking, critical thinking and innovative thinking- (Ramírez-Montoya et al., 2024) with the evidence obtained through exploratory factor analysis, confirmatory factor analysis and the presentation of the instrument's percentiles and scales.

Method

Research Design

This research focuses on seeking and testing the qualities of the eComplexity instrument (Castillo-Martínez and Ramírez-Montoya, 2024) to measure college students' perceptions of their levels of mastery of complex reasoning competence, using (as mentioned by Furr and Bacharach, 2014) a psychometric and instrumental approach to test validity and reliability.

Population and sample

This study included 2205 university students (male: 921; female:1266; “I prefer not to answer”: 18) from many countries, mainly México, España, Chile, Ecuador y Panamá.

Data Collection Tools

The instrument eComplexity initially had 25 items, categorized into four types of thinking (systemic, scientific, critical and innovative). The instrument was submitted for validation by experts, who evaluated the questionnaire according to the criteria of clarity, coherence and relevance. Subsequently the application was self-directed and could be applied both individually and in groups. Each item was answered with a Likert-type scale, with options of (1) No agreement, (2) Little agreement, (3) Neither agree nor disagree, (4) Agreed, (5) Very much in agreement.

Data Collection

The eComplexity questionnaire was applied through a Google form during the period from January 24, 2022 to September 24, 2023 to university students from different countries in Latin America and Spain.

Of the various ways to check content-based validity evidence, the most important is expert support in the subject matter to be studied. This validity evidence tests the degree to which the content aligns with the objectives of the test (Sireci and Faulkner-Bond, 2014). In a previous study, 13 specialists were consulted who averaged 20 years of experience in the field. They considered the form's coherence, relevance, and clarity and provided valuable comments to refine the items. The profile of the experts is shown in Table 1.

Table 1
Profile of experts

Expert	Field of expertise	Current activity	Years of experience
1	Educational innovation	Research Professor	35
2	Linguistics	University professor	15
3	Materials chemistry	Research Professor	20
4	User behaviour	Assistant Research Professor	15
5	Design for Inclusion and Equity. Design and qualitative research methodologies.	Assistant Professor. Department of Design	20
6	Teaching, research in computer science areas	Professor	26
7	Teaching	Plant teacher	25
8	Educational innovation, Adult Education, Open education	Researcher	15
9	Telecommunications, Robotics	Professor	20
10	Educational management and innovation	Research Professor/Postgraduate Program Director	15
11	Statistics, Mathematics, Finance, Decision-Making Modeling	Associate Dean of Research	26
12	Academic	Associate Dean of Research	9
13	University education in science	Science professor	13

The previous study (Ramírez-Montoya et al., 2024) detailed the expert profile and discussed the results of the in-depth expert validation. The averages obtained for the criteria evaluated by the experts were: Clarity: 3.31, which is equivalent to 82.7% of the scale from 1 to 4; Coherence: 3.38, representing 84.5%, and Relevance: 3.54, which is equivalent to 88.5% of the scale from 1 to 4; thus, the criteria were rated at a high level (between 3 and 4).

From the ethical point of view, the necessary requirements were met, through an informed consent, where the objective and the confidential treatment of personal data were explained (Nijhawan, 2013).

Data Analysis

This section presents the different procedures followed to test the efficacy of the eComplexity instrument for the validity and reliability of the evidence. There were different proposals to consider how the analysis of the psychometric properties should be carried out from the different authors' perspectives, so the most relevant ones were considered. Since this research was conducted from the fields of education and psychology, we decided to follow the guidelines recommended by the American Educational Research Association (AERA), American Psychological Association (APA), and National Council on Measurement in Education (NCME) (2014), which describe different types of validity and reliability evidence. Of these, we chose content-based and internal structure-based approaches. For reliability, we chose to use tests of internal consistency (Alpha, Omega, and Lambda).

One of the most critical and sophisticated validity evidence tests is Factor Analysis. Factor analysis is a tool used in the design or validation of psychometric tests (Kaplan & Saccuszzo, 2017). This tool helps in the construction of instruments (Brown, 2015). Factor analysis seeks to determine the variance of the test and the distribution of the items in the minimum number of factors that can explain the test.

Using a significant amount of validity evidence helps achieve a more solid test. There are two factor analysis processes: exploratory factor analysis (EFA) (Lloret-Segura et al., 2014), which helps to determine the number of factors or dimensions that explain the test, and confirmatory factor analysis (CFA) (Keith, 2019), which helps to check the instrument's fit. Both procedures were used in the present research.

Reliability is the accuracy with which an instrument consistently measures participants on a specific variable (Osborne, 2012). There are three types of reliability: test-retest (which helps to check whether the scores are stable over time), parallel testing (applying different versions of the same instrument), and internal consistency (determining whether the items are coherent and congruent with each other) (Abad, 2011). In this research, internal consistency reliability was determined by considering Cronbach's Alpha, McDonald's Omega, and Guttman's Lambda (Table 4).

After checking both the validity and reliability of the instrument, it was necessary to know the centiles and levels that identify the value of the direct scores of the instrument. This process is called bareming, and the designed instruments need to be replicable (Abad, 2011) by knowing the score of the evaluated person and assigning a level. Accordingly, five levels were established, as shown in the results section.

The programs used for the different reliability analyses were the SPSS v. 26 software for the items analyses and the AFE and, finally, the support of JASP for the CFA.

Findings

To facilitate the understanding of the processes that have been carried out, an integrating table (Table 2) is presented that shows the main results, which are subsequently detailed in a more specific manner.

Table 2

Integrated results of psychometric processes

Name	Procedure	Result
Validity evidence based on test content	The instrument was evaluated with the help of 13 expert researchers.	Approval greater than 83%.
Validity evidence based on the internal structure	Exploratory Factor Analysis	KMO: .96 Bartlett’s test: <.05
	Confirmatory Factor Analysis	X2/gl: 7.07; CFI: .918; TLI: .909; GFI: .854; RMSEA: .052; SRMR: .039
Reliability	Internal consistency (ω)	D1: .74; D2: .85; D3: .77; D4: .83; Total: .93

Grading scale	Use of percentiles to obtain the levels	Very low: less than 75 Low: 76 - 91 Medium: 92- 102 High: 103-114 Very high: more than 115

Note: KMO: Kaiser-Mayer-Olkin test; X2: Chi-Square; df: degrees of freedom; CFI: comparative fit index; TLI: Tucker-Lewis index; GFI: goodness-of-fit index; RMSEA: root mean squared error of approximation; SRMR: standardized root mean squared residual; D1: Systemic thinking; D2: Scientific thinking; D3: Critical thinking; D4: Innovational thinking

There are several processes that help to review the psychometric properties of an instrument (AERA et al., 2014), many of them have been captured in Table 2, presented in an integrated and innovative way that helps readers a quick review of the results.

Evidence of validity based on the internal structure

Regarding the exploratory factor analysis (EFA), we obtained a Kaiser-Meyer-Olkin index (KMO) > .80 and a significance of p< .05. These data were very indicative of internal structure since indices greater than .80 are satisfactory (Kaiser, 1979; Lloret-Segura et al., 2014) for a good AFE, where the factor loadings exceed 52% variance across three factors (with a correlation greater than .40).

Subsequently, we proceeded to perform the CFA. This process is a type of structural equation modeling (SEM) that examines the hypothesized relationships between the dimensions and indicators with the latent variables that the indicators are intended to measure (Bollen, 1989; Brown, 2006; Kline, 2010). The Confirmatory Factor Analysis (CFA) for this study is shown in Table 3.

Table 3

Confirmatory Factor Analysis (CFA)

Adjustment indexes	**Model (25 items)**
Parsimony (x^2 /gl)	7.07
CFI	.918
TLI	.909
GFI	.854
RMSEA	.052
SRMR	.039

With the CFA it was possible to detect that some indices could be adjusted to increase the validity of the internal structure of the test (Bentler and Yuan, 1999); therefore, we revised the comparative fit index (CFI), the Tucker-Lewis index (TLI) and the goodness-of-fit index (GFI), which are sought to be greater than .80. We also obtained the root mean squared error of approximation (RMSEA) and the standardized root mean squared residual (SRMR), which are very close to the 5% confidence value that corroborates the validity of the indexes. The analysis allowed us to identify that the items that contributed the least to the CFA were items 8, 22, 23 and 25. However, it was not crucial to eliminate them, since even if the 25 items were kept, the validity of the instrument remained high.

Reliability of the instrument

The results of reliability of the instrument are shown in Table 4 through Cronbach's Alpha, McDonald's Omega, and Guttman's Lambda.

Table 4

Reliability of the instrument

Dimension	McDonald's ω	Cronbach's α	Guttman's λ2	Guttman's λ6
Systemic thinking	.745	.736	.820	.816
Scientific thinking	.858	.857	.832	.828
Critical thinking	.777	.775	.763	.718
Innovative thinking	.836	.835	.842	.818
Reasoning for complexity	.932	.931	.933	.936

Table 4 shows the reliability of the dimensions and the variable. The dimensions of scientific thinking and innovational thinking have values above .80, which indicate reliability. For the variable as a whole, the values are above .91, thus demonstrating high reliability. The dimensions systemic thinking and critical thinking are above .74, acceptable for the instrument's reliability. Therefore, tests for internal consistency proved the instrument's reliability.

Percentile norms and scales of the instrument

Five levels were established through the bareming process, which is shown in Table 5.

Table 5
Bareming process

Percentile	Systemic thinking	Scientific thinking	Critical thinking	Innovative thinking	Reasoning for complexity	Level
5	18	15	16	14	69	Very low
10	20	17	18	16	75	
15	21	18	20	18	80	
20	22	20			83	Low
25		21	21	20	86	
30	23	22	22	20	88	
35	24	23	23	21	91	
40		24	24	22	93	
45		25		23	95	
50	25	26	24	24	98	Average
55		27			99	
60	26	28	25		100	
65					102	
70	27		26	25	104	
75		29	27	26	107	High
80	28	31		27	110	
85	29	32	28	28	114	
90		33	29	29	118	Very High
95	30	35	30	30	123	

	Systemic thinking	Scientific thinking	Critical thinking	Innovational thinking	Reasoning for complexity	
Very low	20 or less	17 or less	18 or less	16 or less	75 or less	Very low
Low	21 - 23	18 - 22	19 - 22	17 - 20	76 - 88	Low
Average	24 - 26	23 - 28	23 - 25	21 - 24	89 - 102	Average
High	27 - 29	29 - 32	26 - 28	25 - 28	103 - 114	High
Very High	30 +	33 +	29 +	29 +	118 +	Very High

Table 5 shows both the percentiles and the instrument's scales, both for the dimensions and the total scale. For example, if a student obtains a direct score of 95 on the total scale, it means that he/she has an average level, and a score of 18 in critical thinking indicates a very low level.

The dimensions and items that make up the eComplexity instrument are listed in Table 6. It is important to mention that in addition to the expert validation and statistical validation shown in the research through the analysis of the psychometric properties, the experts' observations were also taken into consideration to improve the clarity in the wording of some items.

Table 6

Dimensions that make up the eComplexity instrument

	Category	No	Ítem
Reasoning for complexity	Systemic thinking		**Knowledge**
		1	I have the ability to find associations between variables, conditions and constraints in a project.
		2	I identify data from my discipline and other areas that contribute to solving problems.
			Skills
		3	I participate in projects that need to be solved using inter/multidisciplinary perspectives.
		4	I organize information to solve problems.
			Attitudes or values
		5	I like to learn different perspectives on a problem.
		6	I lean towards strategies to understand the parts and the whole of a problem.
	Scientific thinking		**Knowledge**
		7	I have the ability to identify the essential components of a problem to formulate a research question.
		8	I know the structure and formats for research reports used in my area or discipline.
		9	I identify the structure of a research article used in my area or discipline.
			Skills
		10	I identify the elements to formulate a research question.
		11	I design research instruments consistent with the research method used.
		12	I formulate and test research hypotheses.
			Attitudes or values
		13	I am inclined to use scientific data to analyze research problems.
	Critical thinking		**Knowledge**
		14	I have the ability to critically analyze problems from different perspectives.
		15	I identify the basis of my own and others' judgments in order to recognize false arguments.
			Skills
		16	I self-evaluate the level of progress and achievement of my goals to make the necessary adjustments.
		17	I use reasoning based on scientific knowledge to make judgments about a problem.
			Attitudes or values
		18	I make sure to review the ethical guidelines of the projects in which I participate.
		19	I appreciate criticism in the development of projects in order to improve them.
	Innovative thinking		**Knowledge**
		20	I know the criteria to determine a problem.
		21	I have the ability to identify variables, from various disciplines, that can help answer questions.
			Skills
		22	I apply innovative solutions to diverse problems.
		23	I solve problems by interpreting data from different disciplines.
		24	I analyze research problems contemplating the context to create solutions.
			Attitudes or values
		25	I tend to critically and innovatively evaluate the solutions derived from a problem.

Discussion, Conclusion and Implications

A reasoning for complexity instrument encompassing systemic, scientific, critical and innovative thinking can support decision-making for integrated training that prepares students to face current challenges and contributes to educational research. Validating internal structure is useful when designing and implementing an instrument. Regarding the exploratory factor analysis (EFA), a KMO factor > .80 and a significance of $p < .05$ were obtained, which are optimal to propose a good EFA. Exploratory factor analysis makes it possible to identify adequate internal structure in an instrument (Lloret-Segura et al., 2014). Although the PFA allowed obtaining optimal data, it was considered essential to have further support through confirmatory factor analysis (CFA), detailed below.

The confirmatory factor analysis (CFA) allows corroborating whether the dimensions considered for an instrument and its items are adequate. Through confirmatory factor analysis (Table 3), it was possible to identify that both the residual values and the fit values corroborate the internal structure validity of the instrument. Confirmatory factor analysis allows adjusting an instrument until it is optimized for its application (Keith, 2019). In this research, CFA made it possible to distinguish that by retaining the 25 items it was possible to maintain a high validity for the internal structure of the instrument.

Another element to be considered is the instrument's reliability. When calculating the Cronbach's Alpha, McDonald's Omega, and Guttman's Lambda coefficients, we could identify global and individual item optimal values for each dimension, especially for the instrument as a whole, since values in a range of 0.736 to 0.936 were obtained for the various coefficients, as seen in Table 4. Reliability is the precision with which an instrument consistently measures participants on a specific variable (Osborn, 2000). The eComplexity instrument is therefore shown to have an adequate level of reliability.

The norms and percentile scales computed in this study are valuable for upcoming researchers and practitioners who wish to apply the instrument. By knowing the percentiles of the eComplexity instrument, it was possible to visualize the five scale levels from Very Low to Very High (Table 5). The scales were stipulated for the dimensions and the total variables. This procedure is vital when modeling instruments (Abad, 2011). Therefore, the instrument presents levels to be used according to the direct level scores of the persons assessed.

A reasoning for complexity instrument, such as eComplexity, which contemplates the four dimensions of systemic thinking, scientific thinking, critical thinking and innovative thinking can support educational institutions interested in training to solve the challenges of today's society. The value of this instrument is precisely the integration of the four dimensions of thinking so that the results of each dimension can be of value to trainers, managers and decision-makers.

The results showed that the evidence of content-based validity through expert evaluation was favorable for the questionnaire because the ratings for the criteria of Clarity, Coherence, and Relevance were above 3, which placed them at a high level. Based on the evidence of the validity of the internal structure, we proved that the scale adapted to four dimensions (Systemic Thinking, Scientific Thinking, Critical Thinking and Innovative thinking) by testing through Confirmatory Factor Analysis (CFA).

Therefore, the eComplexity instrument is valid and reliable and can be applied in different university contexts of interest, knowing students' perception regarding their mastery levels the complex reasoning competency. It can be stated as an opportunity that the instrument can be applied in different institutions and languages to expand its potential. This paper is an invitation to continue contributing instruments that support training and educational research.

Acknowledgments

The authors acknowledge Tecnologico de Monterrey's financial support through the "Challenge-Based Research Funding Program 2022," Project ID # I005 - IFE001 - C2-T3 – T. The authors acknowledge the financial and technical support of Writing Lab, Institute for the Future of Education, Tecnologico de Monterrey, Mexico, in the production of this work.

References

Abad, F. J. (2011). *Medición en ciencias sociales y de la salud [Measurement in health and social sciences]*. Síntesis

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

Amelink, C. T., Watford, B. A., Scales, G. R., Mohammadi-Aragh, M. J., & Farquhar-Caddell, D. (2013). Defining and measuring innovative thinking among engineering undergraduates. ASEE Annual Conference and Exposition, Conference Proceedings. https://doi.org/10.18260/1-2--19380

Bentler, P. M., & Yuan, K. H. (1999). Structural equation modeling with small samples: Test statistics.

Multivariate Behavioral Research, 34, 181-197.

Bollen, K. A. (1989). A new incremental fit index for general structural equation models. *Sociological methods & research*, 17(3), 303-316.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research* (1a ed.). Guilford publications.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2a ed.). Guilford publications.

Brown, T. A., & Moore, M. T. (2012). Confirmatory factor analysis. *Handbook of structural equation modeling*, 361-379.

Castillo-Martínez, I. M., Ramírez-Montoya, M. S. (2022). eComplexity: Medición de la percepción de estudiantes de educación superior acerca de su competencia de razonamiento para la complejidad. https://hdl.handle.net/11285/643622

Douglas, E. (2006). Critical thinking skills of engineering students: Undergraduate vs. Graduate students. ASEE Annual Conference and Exposition, Conference Proceedings. https://doi.org/10.18260/1-2--101

Furr, R. M., & Bacharach, V. (2014). Test dimensionality and factor analysis. Psychometrics: an introduction, 71-98.

Hernández Sampieri, R., & Mendoza Torres, C. P., 2018. *Metodología de la investigación las rutas cuantitativa, cualitativa y mixta [Research methodology: quantitative, qualitative and mixed routes]*. México: McGraw-Hill.

Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indices in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55

IGI Global. (2021). What is Complex Thinking? Chapter Flipped Classroom, Cognitive Flexibility Theory, and Cognitive Operators of Complexity: Developing TPACK in Pre-Service Teacher Education. *Handbook of Research on TPACK in the Digital Age*. https://www.igi-global.com/dictionary/complex-thinking/66745.

IGI Global. (2021). What is Innovative Thinking? Chapter Teaching Creativity Skills to Foster Social-Emotional Learning. *Aligning Social-Emotional and 21st Century Learning in the Classroom: Emerging Research and Opportunities*. https://www.igi-global.com/dictionary/innovative-thinking/71532.

Izvorska, D. (2016). Educational Researcher A Model for Development of Students' Professional Competency in Technical Universities. *Educational Researcher*, 45, 961–974. https://www.researchgate.net/publication/314062909

Jaaron, A., & Backhouse, C. (2018). Operationalization of service innovation: a systems thinking approach. *The Service Industries Journal*, 38, 561–583. https://doi.org/10.1080/02642069.2017.1411480

Kaiser, H. F., & Cerny, B. A. (1979). Factor analysis of the image correlation matrix. Educational and psychological measurement, 39(4), 711-714.

Kaplan, R. M., & Saccuzzo, D. P. (2017). Psychological testing: Principles, applications, and issues. Cengage Learning.

Keith, T. Z. (2019). *Multiple regression and beyond: An introduction to multiple regression and structural equation modeling*. Routledge.

Kline, R. B. (2006). Structural equation modeling.

Kline, R.B. (2010). Principles and Practice of Structural Equation Modeling. The Guilford Press.

Koerber, S., Mayer, D., Osterhaus, C., Schwippert, K., & Beate, S. (2015). The Development of Scientific Thinking in Elementary School: A Comprehensive Inventory. *Child Development*, 86(1), 327–336. https://doi.org/10.1111/cdev.12298

Lord, F. M., & Novick, M. R. (2008). *Statistical theories of mental test scores*. IAP.

Lloret-Segura, S., Ferreres-Traver, A., Hernandez-Baeza, A., & Tomas-Marco, I. (2014). Exploratory item factor analysis: A practical guide revised and updated. *Anales de Psicología*, 30(3), 1151-1169.

Morin, E. (2003). *A Cabeça Bem-Feita* (8a ed.). Bertrand Brasil.

Morin, E. (2006). *Educar en la era planetaria [Educating in the planetary era]* (1a ed.). Editorial Gedisa

Morin, E. (2015). *Enseñar a vivir[Teaching to live]* (1a ed.). Ediciones Nueva Visión

Morin, E. (2021). Complexidade em diálogos com Edgar Morin[Complexity in dialogues with Edgar Morin]. *Revista Famecos*, 28, 1–5.

Osborne, J. (2012). *Does Reliability Matter?: Debunking the Myth of Perfect Measurement*. SAGE Publications

Powell, N., & Harmon, B. (2016). *Course-Embedded Undergraduate Research Experiences: The Power of Strategic Course Design*. In ACS Symposium Series (pp. 119–136). https://doi.org/10.1021/bk-2016-1231.ch007

Safari, A., Andersson, G., and Kellogg, S. (2019). Understanding Differences in Student Learning Outcomes between the U.S. and Norwegian Educational Systems. Proceedings - Frontiers in Education Conference, FIE, 2019-Octob, 1–5. https://doi.org/10.1109/FIE43999.2019.9028689

Sellars, M., Fakirmohammad, R., Bui, L., Fishetti, J., Niyozov, S., Reynolds, R., Thapliyal, N., Liu-Smith, Y. L., & Ali, N. (2018). Conversations on critical thinking: Can critical thinking find its way forward as the skillset and mindset of the century? *Education Sciences*, 8(4). https://doi.org/10.3390/educsci8040205

Sireci, S., & Faulkner-Bond, M. (2014). Validity evidence based on test content. *Psicothema*, 100-107. 10.7334/psicothema2013.256

Songer, N. B., Kelcey, B., & Gotwals, A. W. (2009). How and when does complex reasoning occur? Empirically driven development of a learning progression focused on complex reasoning about biodiversity. *Journal of Research in Science Teaching*, 46(6), 610–631. https://doi.org/10.1002/tea.20313

Straková, Z., & Cimermanová, I. (2018). Critical thinking development-a necessary step in higher education transformation towards sustainability. *Sustainability* (Switzerland), 10, 1–18. https://doi.org/10.3390/su10103366

Suryansyah, A., Kastolani, W., & Somantri, L. (2021). Scientific thinking skills in solving global warming problems. IOP Conference Series: Earth and Environmental Science, 683(1). https://doi.org/10.1088/1755-1315/683/1/012025

Tecnológico de Monterrey. (2019). Razonamiento para la complejidad. In Competencias Transversales. Una visión desde el modelo educativo TEC21. *Documento guía para el docente de educación superior [Higher Education Teacher's Guide]* (pp. 62–76). Tecnológico de Monterrey

Zimmerman, C., & Croker, S. (2014). A prospective cognition analysis of scientific thinking and the implications for teaching and learning science. *J. Cogn. Educ. Psychol*, 13, 245–257. https://doi.org/10.1891/1945-8959.13. 2.245