STUDIA UBB PSYCHOL.-PAED., LXIX, 1, 2024, p. 5-26
(Recommended Citation)
DOI:10.24193/subbpsyped.2024.1.01

The comparison of several factorial structures of the Cornell Critical Thinking Test Level Z

Andrea BARTA[1,*], Borbála TAMÁS[1], Tünde PÓKA[1,2]

ABSTRACT. The Cornell Critical Thinking Test (CCTT) level Z is a broadly applied instrument for the measurement of undergraduate, gifted students' and adults' critical thinking abilities, however, there is a lack of studies in the literature investigating its factorial structure. The test developers emphasize the interpretation of the test results as a general critical thinking factor, including different cognitive abilities, but also highlight the overlap between the items of the subfactors. The aim of the study is the investigation of internal consistency, the comparison of different factorial structures (unidimensional, correlated, and hierarchical models), and gender invariance testing of the CCTT level Z. Hungarian-speaking undergraduate students participated in the study (N = 825). For the confirmatory factor analysis (CFA) the Mplus version 8.7, with weighted least squares mean and variance adjusted estimation was used. Measurement invariance of the test across genders was analyzed using Muthén's two-step procedure for dichotomous data. The results indicated that the correlated four-factor and second-order structures of the test exceeded the acceptable model fit criteria. Post hoc inspection was conducted on the second-order four-factor model, indicating a shortened 22-item version of the test (with a general critical thinking factor, including four subfactors: deduction, meaning and fallacies, induction, assumption identification) with excellent fit indices [$\chi 2$ (203) = 259.309, CFI = .967, TLI = .963, RMSEA = .018, SRMR = .056]. Configural and scalar invariance of the abbreviated Hungarian version of the test across genders were confirmed.

Keywords: critical thinking, undergraduate students, confirmatory factor analysis, measurement invariance

[1] Department of Applied Psychology, Faculty of Psychology and Educational Sciences, Babeş-Bolyai University, Cluj-Napoca, Romania
[2] Evidence Based Psychological Assessment and Interventions Doctoral School, Babeş-Bolyai University, Cluj-Napoca, Romania
* Corresponding author: andrea.barta@ubbcluj.ro

©2024 STUDIA UBB PSYCHOL.-PAED. Published by Babeş-Bolyai University.
This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

1. INTRODUCTION

Critical thinking is a higher-order cognitive ability that also includes critical thinking disposition, personality, and motivational and cognitive characteristics that contribute to the application of critical thinking abilities (Ennis, 2018; Facione, 1990a; Haber, 2020; Halpern & Sternberg, 2019). There is no consensus in the literature regarding the definition of critical thinking. Watson and Glaser (1980) emphasize five cognitive skills within critical thinking: interpretation, recognition of assumptions, deduction, inference-making, and evaluation of arguments. Furthermore, they also highlight the role of the improvement of critical thinking for the increase of academic success (Aiyub et al., 2021). Facione (1990a) differentiated six main critical thinking skills: interpretation, analysis, evaluation, conclusion, explanation, and self-regulation. Ennis (2018) grouped critical thinking abilities into four main categories (problem clarification abilities, advanced clarification abilities, decision-making abilities, and inference abilities), which include several specific subskills. Halpern (1997) defined critical thinking as the totality of cognitive skills and strategies applied in problem-solving, drawing conclusions, and decision-making. The aim of the APA Delphi study was the unified definition of critical thinking. As a result of this study, the researchers concluded that analysis, evaluation, inference, interpretation, explanation, and self-regulation are the basic critical thinking skills (Facione, 1990b).

Critical thinking, as a 21st-century ability, was examined in several psychological areas. Educational psychology research concentrates on the role of critical thinking in learning, teaching, and learning achievement, respectively on its improvement in the case of students as well as in the case of teachers (Bezanilla et al., 2019; Cruz et al., 2019; D'Alessio et al., 2019; Janssen et al., 2019; Renatovna, & Renatovna, 2021; Yuan et al., 2022). Cognitive psychology studies focus on the association between critical thinking and other cognitive processes and also investigate the role of critical thinking in complex cognitive activities, like problem-solving (Işıklar, & Abalı-Öztürk, 2022; Muhammad Raflee, & Halim, 2021; Sholihah, & Lastariwati, 2020; Song et al., 2022) and metacognition (Boran, & Karakuş, 2022; Danial et al., 2018; Deliligka, & Calfoglou, 2022; Jin, & Ji, 2021; Marthaliakirana et al., 2022; Murtadho, 2021). In clinical psychology, it is mainly the role of the maladaptive form of critical thinking, and self-criticism that is studied in several mental disorders (Harman, & Lee, 2009; Moreira, & Canavarro, 2018; Moroz, & Dunkley, 2019; Wakelin et al., 2022). In personality psychology, the relationships between critical thinking dispositions and personality traits (Eshmirzaeva, 2020; Fitriana et al., 2018; Toker, & Akbay, 2022), as well as effective personality (Merma-Molina et al., 2022) were measured.

The level of development of critical thinking has an effect on individual achievement, success, and prosperity in every stage of life (Evans, 2020; Franco et al., 2018; van der Zanden et al., 2019; Živković, 2016) because critical thinking abilities are cognitive processes that are indispensable instruments of learning, problem-solving, and innovation (Culver et al., 2019; Ericson, 2022; Mujanah et al., 2022; Muhammad Raflee, & Halim, 2021; Sari, & Wardhani, 2020; Song et al., 2022). Several studies demonstrated positive associations between critical thinking skills and learning performance (Akpur, 2020; Fatmawati et al., 2019; Ghanizadeh, 2017; Ibrahim et al. 2021; Ng et al., 2022; Shahzadi et al., 2020). After graduating, critical thinking also promotes fulfillment in the job market and career success (Crosta, & Banda, 2022; Jebreen, & Nabot, 2021; Okolie et al., 2022; Saleh, 2019; Teng et al., 2019).

Due to the comprehensive applicability of critical thinking, the assessment and development of this cognitive process have a critical role. Reliable measurement of critical thinking abilities can provide relevant information about the level of development of a student's critical thinking skills and which skills require further development to achieve successful academic results. The Cornell Critical Thinking Test Level Z is one of the most commonly used assessment tools in the academic context to measure students' critical thinking abilities. However, there is no prior research in the literature regarding the test's factorial structure. Furthermore, there is no previously validated tool for measuring critical thinking among Hungarian-speaking university students.

1.1. The measurement of critical thinking

Several instruments have been developed for the measurement of critical thinking abilities and dispositions. For the assessment of critical thinking, standardized tests are applied most frequently (Gunawardena, & Wilson, 2021; Imperio et al., 2020). There are overlaps between the cognitive abilities that are measured with different tests, but some assessed abilities differ in terms of the theoretical models of the tests, so scientists defined the term of critical thinking differently and emphasized different cognitive processes, or activities within critical thinking. A further difference between the tests is that they could require recognition or recall memory. Tests and exercises requiring recall memory have short answers, or short essays as answers, while the tests calling for recognition memory have a forced-choice question format (Butler et al., 2012).

One of the most frequently applied instruments for the assessment of critical thinking is the Cornell Critical Thinking Test (Alias et al., 2022; Ennis et al., 2005), which is based on the Cornell/Illinois model, and measures the deduction, induction, credibility, observation, and assumption identification abilities within critical thinking (Ennis et al., 2005; Imperio et al., 2020). Another very widely used, recognition-requiring tool for the assessment of adults' critical thinking, is the California Critical Thinking Skills Test, developed by Facione (1990b) based on the conclusions of the Delphi report. The test measures critical thinking skills (interpretation, analysis, evaluation, inference, explanation, except for self-regulated learning) approved by the 46 experts contributing to the Delphi report (Alias et al., 2022; Butler et al., 2012; Facione, 1990b).

The Ennis-Weir Critical Thinking Essay Test (Ennis, & Weir, 1985) contains open-ended questions. The test taker reads a letter relating a fictive problem-solving situation including a list of arguments, then the responder needs to express their point of view supporting it with arguments in an essay format. The responder needs to react, give an answer in every paragraph, and then needs to explain their viewpoint regarding the whole text. They receive points for identifying the strengths and weaknesses of the read arguments, and respectively for the arguments that support their decisions. Compared with the forced-choice format, the test assesses not only the evaluative function of critical thinking but also the creative, productive characteristics of it, thereby it is a more holistic instrument for the assessment of critical thinking. However, the subjectivity resulting from the evaluation of the answers may lead to biases (Alias et al., 2022; Butler et al., 2012; Ennis, & Weir, 1985; Werner, 1991). Another test requiring recall memory for the measurement of critical thinking is the Watson and Glaser Critical Thinking Appraisal (Watson, & Glaser, 1980) which examines critical thinking during problem-solving. Initially, the test assessed five abilities within critical thinking (recognition of assumptions, evaluation of arguments, deduction, inference making, interpretation), later they operationalized the intercorrelated interpretation, inference, and deduction abilities as one overall factor, which was called conclusions. The abilities of recognition of assumptions and evaluation of arguments were still treated as separate factors (Aiyub et al., 2021; Alias et al., 2022; Butler et al., 2012; Watson, & Glaser, 2010).

Contrary to the Cornell Critical Thinking Test and the California Critical Thinking Skills Test, which have forced-choice questions, respectively the Ennis-Weir Critical Thinking Essay Test and the Watson-Glaser Critical Thinking Appraisal, which require recall memory, the Halpern Critical Thinking Assessment combines recall and recognition memory, contains open-ended and also forced-choice questions regarding the read situations (Butler et al., 2012; Gunawardena, & Wilson, 2021; Alias et al., 2022). The test measures the following critical thinking abilities: verbal reasoning, argument analysis, skills in thinking (hypothesis testing), using likelihood and uncertainty, decision-making, and problem-solving (Halpern, 1997; Butler et al., 2012).

1.2. The internal structure of Cornell Critical Thinking Test Level Z

The Cornell Critical Thinking Test rests on the Cornell/Illinois model (Gunawardena & Wilson, 2021; Imperio et al., 2020; Ling & Loh, 2020). Ennis differentiated three cognitive abilities within critical thinking: induction, deduction, and value judgment, and four methods applied during these cognitive processes: assumptions, inferences, observations, and statements (Ennis, 2018; Ennis et al., 2005; Ling & Loh, 2020). The model based on the interaction of these cognitive abilities and methods provided the basis of the Cornell Critical Thinking Test (Ennis et al., 2005; Ling & Loh, 2020). Two versions of the test were developed for the measurement of critical thinking in two different age groups. The Cornell Critical Thinking Test Level X was developed for the measurement of critical thinking of middle school and high school students (4-12 grades), while the Cornell Critical Thinking Test Level Z is applicable for the measurement of critical thinking abilities of advanced and gifted high school students, college students and adults (Ennis et al., 2005). Both tests have a forced-choice question format and dichotomous items. The Cornell Critical Thinking Test Level Z has 52 items and measures cognitive abilities such as deduction, induction, observation, credibility, meaning, and assumption identification (Gunawardena & Wilson, 2021; Ennis et al., 2005; Verburgh et al., 2013). The CCTT is a domain-general instrument, the results are independent of domain-specific knowledge.

The CCTT Level Z is a widely used instrument for the measurement of students' critical thinking in educational, cognitive psychology studies (Bataineh, & Zghoul, 2006; Beavers et al., 2017; Frost et al., 2019; Heidari, 2020; Kusumoto, 2018; Saud, 2020). However, there is a lack of empirical studies in the literature investigating the factorial structure of the CCTT Level Z. The instrument contains seven subtests: deduction, meaning and fallacies, observation and credibility of sources, induction: hypothesis testing, induction: planning experiments, definition, and assumption identification. As seen in the enumeration above, the induction ability is composed of two subscales. The test measures the application of induction in hypothesis testing and planning experiments. Similarly, the test contains two sections for the measurement of assumption identification ability (Ennis et al., 2005). The test authors emphasize the interdependency and overlap of the measured cognitive abilities. Due to the reduced number of items assessing the separate cognitive abilities, the test authors suggest taking into consideration the total score for concluding individual differences in critical thinking. In the manual of the Cornell Critical Thinking Test, the reliability results of the instrument regarding undergraduate and graduate student samples are reported. The Kuder-Richardson reliabilities ranged between .50 and .76, and the split-half reliabilities were between .49 and .80 (Ennis et al., 2005; Verburgh et al., 2013).

1.3. The aims of the study

The main aim of the study was the translation into Hungarian language of the first instrument that measures the critical thinking abilities of high school students, gifted students, and adults. Due to the lack of earlier studies investigating the factorial structure of the Cornell Critical Thinking Test Level Z, based on the recommendations of the authors of the test (Ennis et al., 2005), besides the unidimensional model, our aim was to test and compare several correlated and hierarchical factorial models (two-factor models: deductive reasoning, inductive reasoning; three-factor models: deduction, induction, meaning and fallacies; four-factor models: deduction, induction, meaning and fallacies, assumption identification) translated into Hungarian language. Our further aim was the invariance testing of the CCTT across genders.

2. METHODS

2.1. Participants

825 Hungarian-speaking undergraduate students from Babeș-Bolyai University participated in the study. The participants had given their written consent to voluntary participation in the research and the anonymous use of their data. There were no multiple outliers identified, based on the Mahalanobis distance. 78.3% of the participants were females, with a mean age of 21.76 years (SD = 7.12). The youngest person was 18, and the oldest was 64 years old. The majority of participants were first-year (85%), 13.9% of them were third year and only 1.1% of the participants were second-year undergraduate students. 86.7% of the participants were full-time students, and 42.2% of them studied at the Faculty of Psychology and Educational Sciences. The detailed Faculty distribution of the participants is presented in Table 1.

Table 1. *Faculty-based distribution of the participants*

Faculty	N	%
Psychology and Educational Sciences	348	42.2
Geography	70	8.5
Mathematics and Computer Science	68	8.2
Physical Education and Sport	62	7.5
Letters / Humanities	58	7.0
Biology and Geology	40	4.8
Reformed Theology and Music	32	3.9

Faculty	N	%
Sociology and Social Work	30	3.6
History and Philosophy	29	3.5
Chemistry and Chemical Engineering	21	2.5
Political, Administrative and Communication Sciences	19	2.3
Economics and Business Administration	19	2.3
Physics	9	1.1
Theatre and Film	7	.8
Law	6	.7
Roman Catholic Theology	4	.5
Environmental Science and Engineering	3	.4

2.2. Instrument

The Cornell Critical Thinking Test Level Z was applied for the measurement of students' higher-order thinking abilities (Ennis et al., 2005). The CCTT contains 52 items and seven sections: deduction (1-10), meaning and fallacies (11-21), observation and credibility of sources (22-25), induction (Hypothesis Testing) (26-38), induction (Planning Experiments) (39-42), definition and assumption identification (43-46), assumption identification (47-52). The CCTT is a multiple-choice test, has a forced-choice question format, and dichotomously scored items, the answer for an item can be correct or incorrect. The retranslation method was applied for the translation of the instrument into Hungarian language, executed by two experts. The participation in the study was preceded by the informed consent of the participants, the test was completed in Google Forms.

2.3. Data analysis

For the CFA analysis, we applied the Mplus version 8.7 with weighted least squares, mean, and variance-adjusted estimation (WLSMV; Muthén et al., 1997; Muthén & Muthén, 2021). The following absolute fit indices were used to evaluate model-data fit: Chi-squared, Root Mean Squared Error of Approximation (RMSEA), and Standardized Root Mean Square Residual (SRMR). Some relative fit indices, like the Comparative Fit Index (CFI), and Tucker-Lewis Index (TLI) were also applied. Acceptable model-data fit criteria included CFI > 0.90, TLI > 0.90, RMSEA < 0.05, SRMR < 0.08 (Awang, 2012; Hu & Bentler, 1999; Kline, 2015), while models having CFI and TLI values greater than .95 indicate excellent fit (Hu & Bentler, 1999; Kline, 2015). The relative fit of the alternative factor structures was compared based on the difference between CFI values. Post hoc inspection was conducted on the best-fitting model, and the following item retention criteria were applied: items with significant factor loadings and items with factor loadings ≥ 0.32 (Costello & Osborne, 2005; Leach et al., 2020).

We also tested measurement invariance across genders with Muthén's (2013) two-step procedure for dichotomous data (Leach et al., 2020). We have tested configural and scalar invariance models because metric invariance testing is not allowed for binary variables. The configural model was applied to determine if the four-factor structure existed across the two groups, without applying equality constraints. In the scalar invariance model, the factor loadings and intercepts were constrained as equal across groups (Leach et al., 2020; Muthén, 2013; Muthén & Muthén, 2017). Due to our relatively large sample size and the sensitivity of the absolute $\Delta\chi2$ to sample size, ΔRMSEA and ΔSRMR were used for testing the change in model fit, values < .015 and < .01 indicated no difference between the models (Chen, 2007; Meade et al., 2008).

3. RESULTS

The descriptive statistics of the *CCTT Level Z* items are presented in Table 2. Item means varied substantially with values ranging from .11 (Item 18) to .80 (Item 8), and standard deviations ranged from .30 (Item 18) to .50 (Items 4 and 29). These values reveal that several items were very difficult (e.g., Items 12, 18, 32, 37), while other items were easier (e.g., Items 2, 8, 17, 26, 46).

Table 2. Means and Standard Deviations of CCTT level Z Hungarian version items

Item	*M*	*SD*	Item	*M*	*SD*
1	.45	.49	27	.58	.49
2	.74	.43	28	.72	.45
3	.30	.45	29	.51	.50
4	.50	.50	30	.45	.49
5	.65	.47	31	.27	.44
6	.42	.49	32	.12	.32
7	.40	.49	33	.64	.48
8	.80	.40	34	.64	.47
9	.37	.48	35	.40	.49
10	.73	.44	36	.56	.49
11	.60	.49	37	.15	.36
12	.15	.35	38	.36	.48

Item	*M*	*SD*	Item	*M*	*SD*
13	.30	.45	39	.32	.46
14	.16	.36	40	.46	.49
15	.28	.45	41	.35	.47
16	.28	.44	42	.53	.49
17	.74	.44	43	.39	.48
18	.11	.30	44	.57	.49
19	.25	.43	45	.36	.47
20	.16	.36	46	.79	.40
21	.17	.38	47	.51	.50
22	.31	.46	48	.58	.49
23	.43	.49	49	.72	.44
24	.46	.49	50	.61	.48
25	.36	.48	51	.43	.49
26	.74	.43	52	.27	.44

3.1. The comparison of several factorial structures of the CCTT Level Z in Hungarian

Based on the recommendations of the test authors (Ennis et al., 2005), beside the unidimensional model, several correlated and hierarchical models [(two-factor models: deductive reasoning (Items 1-25, 43-52), inductive reasoning (Items 26-42); three-factor models: deduction (Items 1-10, 43-52), meaning and fallacies (Items 11-21), induction (Items 26-42); four-factor models: deduction (Items 1-10), meaning and fallacies (Items 11-21), induction (Items 26-42), assumptions (43-52)] were also tested. The analysis of the observation and credibility of sources as a discrete factor and the items belonging to this factor were excluded from the CFA analysis due to the insignificant factor loadings of all four items and negative covariances, residual variances with other latent variables. Due to the overparameterized factor structure, negative covariances, and residual variances, the items belonging to the induction-hypothesis testing and induction-planning experiments subscales were treated as the induction factor, similarly, the definition-assumption identification and assumption identification subscales were also treated as a single factor. The results of the tested CFA models are presented in Table 3.

Table 3. Goodness of fit statistics of the tested CFA models

Model	$\chi2$	df	p	RMSEA	90% CI		SRMR	CFI	TLI
					LI	UI			
Unidimensional	1557.576	1267	<.001	.017	.014	.019	.064	.860	.854
Two-factor correlated	1508.969	1266	<.001	.015	.012	.018	.063	.883	.877
Two-factor second-order	1508.969	1266	<.001	.015	.012	.018	.063	.883	.877
Three-factor correlated	1255.214	1071	<.001	.014	.011	.018	.062	.906	.901
Three-factor second-order	1255.214	1071	<.001	.014	.011	.018	.062	.906	.901
Four-factor correlated	1247.415	1068	<.001	.014	.010	.018	.061	.909	.904
Four-factor second-order	1248.971	1070	<.001	.014	.010	.018	.061	.909	.904
Four-factor second-order abbreviated model	259.309	203	.005	.018	.011	.025	.056	.967	.963

The CFA results of the tested models revealed that the four-factor correlated [$\chi2$ (1068) = 1247.415, CFI = .909, TLI = .904, RMSEA = .014, SRMR = .061] and the four-factor second order models [$\chi2$ (1070) = 1248.971, CFI = .909, TLI = .904, RMSEA = .014, SRMR = .061] exceeded the minimum criteria for acceptable model fit on the majority of fit indices (Awang, 2012; Hu & Bentler, 1999; Kline, 2015).

3.2. The abbreviated Hungarian version of the CCTT Level Z

Due to the very similar statistical fit indices of the two four-factor models, for post hoc inspection the second-order four-factor model was chosen based on the theoretical approach of the test developers, who suggest the interpretation of critical thinking as a construct involving subfactors (Ennis et al., 2005). Based on the post hoc inspection of the four-factor second-order model, items loaded insignificantly to the deduction factor (3, 4) were not retained. Similarly, three items were identified (13, 15, 19) that were not loaded significantly to the meaning and fallacies factor, two items with insignificant factor loadings to the induction factor (39, 40), and one item (43) that loaded insignificantly to the assumption identification factor. Within the deduction subfactor five items were identified, (1, 2, 6, 9, 10) within the meaning and fallacies subfactor five items (12, 14, 16, 20, 21), within the induction subfactor five items (31, 35, 38, 41, 42), and within the assumptions subfactor four items (44, 45, 51, 52) with factor loadings < .32. After the exclusion of items with insignificant and lower than .32 factor loadings (Costello, & Osborne, 2005; Leach et al., 2020), the abbreviated version of the test included 22 items, three items (5, 7, 8) loaded significantly to the deduction factor, four items (11, 17, 18, 20) to the meaning and fallacies factor, ten items (26, 27, 28, 29, 30, 32, 33, 34, 36, 37) to the induction factor and five items (46, 47, 48, 49, 50) to the assumption identification factor. The abbreviated 22-item four-factor second-order model indicated excellent fit indices [$\chi 2$ (203) = 259.309, CFI = .967, TLI = .963, RMSEA = .018, SRMR = .056]. The factor loadings of the abbreviated model are presented in Table 4.

Table 4. Standardized factor loadings of the abbreviated Hungarian 22-item four-factor second order model

Item	Factor	Standardized		
		Estimate	SE	RV
5	Deduction	.417	.062	.826
7	Deduction	.332	.060	.890
8	Deduction	.576	.069	.669
11	Meaning and fallacies	.429	.060	.816
17	Meaning and fallacies	.456	.067	.792
18	Meaning and fallacies	-.370	.082	.863
20	Meaning and fallacies	-.341	.074	.884
26	Induction	.485	.050	.765
27	Induction	.567	.048	.678
28	Induction	.668	.045	.554
29	Induction	.403	.050	.837
30	Induction	.578	.048	.666
32	Induction	-.371	.063	.862
33	Induction	.322	.054	.897
34	Induction	.535	.048	.714
36	Induction	.506	.049	.744
37	Induction	-.407	.064	.834
46	Assumption identification	.631	.057	.601
47	Assumption identification	.343	.056	.883
48	Assumption identification	.340	.056	.885
49	Assumption identification	.514	.056	.735
50	Assumption identification	.616	.053	.621

Notes. SE = Standard Error; RV = residual variance

Table 5 presents the descriptive statistics of the abbreviated 22-item four-factor second-order model.

Table 5. Descriptive statistics of the CCTT level Z Hungarian version

Factor	*M*	*SD*	*Min.*	*Max.*
Deduction	1.85	.86	0	3
Meaning and fallacies	1.59	.77	0	4
Induction	5.11	2.00	0	9
Assumption identification	3.21	1.30	0	5
Total critical thinking	11.77	3.23	2	19

The internal consistency of the 22-item Hungarian version of the test was acceptable ($\alpha = .601$).

3.3. Measurement invariance across genders of the Hungarian version of the CCTT level Z

Measurement invariance across genders (males and females) of the 22-item abbreviated version of the Cornell Critical Thinking Test Level Z was tested using configural and scalar models. The models showed acceptable model fit indices for both genders at every level of invariance. The configural model indicated the 22-item four-factor structure for males and females. The scalar model indicated that item loadings onto the factors and item intercepts are similar across gender groups. The changes in fit indices (ΔCFI and ΔRMSEA) indicated no significant differences in relative model fit between the configural and the restrictive model, which confirms the configural and scalar invariance of the instrument across genders. Table 6 presents the fit indices for configural and scalar invariance models across genders for the 22-item solution.

Table 6. Results of Measurement Invariance between genders ($n_{females} = 645$; $n_{males} = 179$)

Model	$\chi2$ (df)	$\Delta\chi2$	CFI	ΔCFI	RMSEA	90% CI	ΔRMSEA
Configural	450.724 (402)*	15.341	.968	.001	.017	.002 - .025	< .001
Scalar	466.065 (416)*		.967		.017	.003 - .025	

Notes. * $p < .05$.

4. DISCUSSION

The main aim of the study was the translation into Hungarian and the investigation of the factorial structure of the CCTT Level Z with confirmatory factor analysis. Due to the lack of studies investigating the factorial structure of the test, we analyzed several factorial models (unidimensional, two-factor correlated, two-factor hierarchical, three-factor correlated, three-factor hierarchical, four-factor correlated and four-factor hierarchical) based on the recommendations of the test authors (Ennis et al., 2005).

The results of the CFAs indicated that the four-factor correlated, and the four-factor hierarchical models exceeded our criteria for acceptable model fit, except for the $\chi 2$ indicator (Awang, 2012; Hu & Bentler, 1999; Kline, 2015). Post hoc inspection was also conducted for the detection of insignificantly loading items to the factors and for the exclusion of items with factor loadings ≥ 0.32 (Costello, & Osborne, 2005; Leach et al., 2020). From the two four-factor models (including the induction, meaning and fallacies, deduction, and assumption identification factors), the second-order model was chosen for post hoc inspection, based on the theoretical framework of the test authors (Ennis et al., 2005), who defined critical thinking as a higher order cognitive construct including different critical thinking abilities. The abbreviated 22-item four-factor second-order Hungarian version of the CCTT exceeded our criteria for excellent model fit (Hu, & Bentler, 1999; Kline, 2015).

The internal consistency of the shortened version of the test is similar to the findings of test authors (Ennis et al., 2005), who examined the CCTT's internal consistency in the case of several undergraduate and graduate samples, revealing Kuder-Richardson reliabilities between .50 and .76. The current internal consistency indicator (.601) exceeds the reliability of the original version of CCTT applied in the study of Verburgh et al. (2013), who found a lower Cronbach alpha indicator (.52) in the case of Belgian Educational Science students.

The results regarding the measurement invariance test of the 22-item Hungarian version of the CCTT indicated configural and scalar invariance across genders, with the two tested models having acceptable fit indices for both the male and female groups. The configural model demonstrated the 22-item four-factor structure for both genders, and the scalar model showed that item loadings and item intercepts are similar across the two gender groups.

Based on our results, the 22-item four-factor second-order version of CCTT is reliably applicable in the case of Hungarian-speaking undergraduate students. Our findings support the approach of the test developers concerning the interpretation of the results (Ennis et al., 2005), indicating the presence of a general factor that can be used for the comparison of the critical thinking of students, adults from different groups; but also highlight the existence of different cognitive abilities (deduction, meaning and fallacies, induction, assumption identification) within critical thinking that can be interpreted separately in the case of Hungarian-speaking students population.

This is the first study that aimed at the comparison of different factorial structures of the CCTT level Z. Similarly, there is a lack of instruments in Hungarian language measuring students' and adults' critical thinking. The results of this study, the translated and validated Hungarian version of the CCTT level Z, are very important for the empirical measurement of critical thinking skills in education, and in different work environments. The empirical measurement of students' critical thinking provides information about their level of thinking skills and promotes the identification of those skills that require further development in the educational context.

4.1. Limitations and future directions

One of the limitations of our study is that the instrument was applied only to Hungarian-speaking undergraduate students, so our results cannot be generalized to Romanian-speaking and other language-speaking students. Our further aim is the application of the test and the investigation of its factorial structure on other undergraduate student samples, including other universities, the translation and validation of the test into Romanian, besides the Hungarian language version to increase the generalizability of the results and the literature regarding the CCTT's internal consistency and its factorial structure.

Another limitation of the study is a consequence of the convenience sampling method, that is the unequal distribution of participants based on their socio-demographic characteristics. The majority of participants were first-year undergraduate females, who studied at the Faculty of Psychology and Educational Sciences, so the social science academic discipline was overrepresented compared to other academic disciplines, like humanities, natural sciences, sports, and arts. In the future, it would be meritorious to apply the cluster sampling method for the equal representation of students based on gender, academic discipline, and academic year. Additionally, our aim is the application of the CCTT to master students, besides undergraduate students.

It is very important to continue to analyze the psychometric properties, validity, and reliability of the test on different samples, due to the reduced number of studies providing empirical evidence regarding these characteristics of the CCTT level Z. The measurement invariance of the test across genders was conducted on an unequally distributed gender sample, the number of female participants was almost four times higher than that of male participants. Future studies should focus on the analysis of the measurement invariance of the test across genders on samples with balanced gender distribution. It would also be worth to assess the test's measurement invariance across other important socio-demographic (e.g. age, parents' educational level) and academic variables (e.g. academic achievement, earlier academic degree, academic discipline) of students that could influence their critical thinking level, and the factorial structure of the test could change across different groups of students.

5. CONCLUSION

The CCTT is an extensively used instrument for the measurement of critical thinking in different psychological areas, but there is a lack of earlier studies in the literature investigating its factorial structure and reliability. Due to the questionable factorial structure of the test and the absence of instruments assessing adults' critical thinking abilities, the aim of the current study was the translation of the test into Hungarian language and the analysis of different factorial structures proposed by the test developers (Ennis et al., 2005). Comparing the unidimensional model with different correlated and hierarchical factorial structures, our results revealed that the four-factor structure of the test, including four higher-order cognitive abilities within the general critical thinking factor, namely the deduction, meaning and fallacies, induction and assumption identification factors, is the most applicable, reliable and valid model for the measurement of the level of critical thinking of undergraduate Hungarian-speaking students. Post hoc inspection of the four-factor second-order structure indicated a 22-item shortened version of the test with excellent fit indices.

Additionally, invariance testing of the instrument across genders was conducted. The results revealed configural and scalar invariance across genders of the 22-item four-factor structure Hungarian version of the CCTT, confirming that the test has similar factorial structure, item loadings, and intercepts in the case of both male and female groups. The internal consistency of the abbreviated test is similar to other results regarding the reliability of the test (Ennis et al., 2005; Verburgh et al., 2013).

Based on the limitations of the study, for future research it is recommended to evaluate the factorial structure and reliability of the instrument on older master students, on samples with balanced gender, academic discipline, and academic year distribution, replacing the convenience sampling method with the cluster one. Similarly, it would be beneficial to translate and validate the test into other languages, its measurement invariance analysis between groups formed based on socio-demographic and academic characteristics of students, to complete the literature with additional results regarding the factorial structure and the reliability of the test.

Statements and declarations. Declaration of competing interest. None.

Contributions. All authors contributed substantially to the study and approved the submitted version.

REFERENCES

Aiyub, Suryadi, D., Fatimah, S., & Kusnandi. (2021). Investigation of Watson-Glaser critical thinking skills of junior high school students in solving mathematical problems. *Journal of Physics: Conference Series*, *1806*(1), 012090. https://doi.org/10.1088/1742-6596/1806/1/012090

Akpur, U. (2020). Critical, Reflective, Creative Thinking and Their Reflections on Academic Achievement. *Thinking Skills and Creativity*, *37*, 100683. https://doi.org/10.1016/j.tsc.2020.100683

Alias, A., Mohtar, L. E., Ayop, S. K., & Rahim, F. R. (2022). A Systematic Review on Instruments to Assess Critical Thinking & Problem-Solving Skills. *EDUCATUM Journal of Science, Mathematics and Technology*, *9*, 38-47. https://doi.org/10.37134/ejsmt.vol9.sp.5.2022

Awang, Z. (2012). *Research methodology and data analysis.* Penerbit Universiti Teknologi MARA Press.

Bataineh, R. F., & Zghoul, L. H. (2006). Jordanian TEFL Graduate Students' Use of Critical Thinking Skills (as Measured by the Cornell Critical Thinking Test, Level Z). *International Journal of Bilingual Education and Bilingualism*, *9*(1), 33–50. https://doi.org/10.1080/13670050608668629

Beavers, E., Orange, A., & Kirkwood, D. (2017). Fostering critical and reflective thinking in an authentic learning situation. *Journal of Early Childhood Teacher Education*, *38*(1), 3–18. https://doi.org/10.1080/10901027.2016.1274693

Bezanilla, M. J., Fernández-Nogueira, D., Poblete, M., & Galindo-Domínguez, H. (2019). Methodologies for teaching-learning critical thinking in higher education: The teacher's view. *Thinking Skills and Creativity*, *33*, 100584. https://doi.org/10.1016/j.tsc.2019.100584

Boran, M., & Karakuş, F. (2022). The Mediator Role of Critical Thinking Disposition in the Relationship between Perceived Problem-Solving Skills and Metacognitive Awareness of Gifted and Talented Students. *Participatory Educational Research*, *9*(1), 61–72. https://doi.org/10.17275/per.22.4.9.1

Butler, H. A., Dwyer, C. P., Hogan, M. J., Franco, A., Rivas, S. F., Saiz, C., & Almeida, L. S. (2012). The Halpern Critical Thinking Assessment and real-world outcomes: Cross-national applications. *Thinking Skills and Creativity*, *7*(2), 112–121. https://doi.org/10.1016/j.tsc.2012.04.001

Chen, F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14,* 464–504. https://doi.org/10.1080/10705510701301834

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, *10*, 1-9.

Crosta, L., & Banda, V. (2022). Are young graduates ready for the job market of the future? The study of the Italian case. *GiLE Journal of Skills Development*, *2*(1), 26–43. https://doi.org/10.52398/gjsd.2022.v2.i1.pp26-43

Cruz, G., Nascimento, M. M., & Dominguez, C. (2019). With a little help from my peers: Professional development of higher education teachers to teach critical thinking. *Revista Lusófona de Educação*, *44*, 141–157. https://doi.org/10.24140/issn.1645-7250.rle44.09

Culver, K. C., Braxton, J., & Pascarella, E. (2019). Does teaching rigorously really enhance undergraduates' intellectual development? The relationship of academic rigor with critical thinking skills and lifelong learning motivations. *Higher Education*, *78*(4), 611–627. https://doi.org/10.1007/s10734-019-00361-z

D'Alessio, F. A., Avolio, B. E., & Charles, V. (2019). Studying the impact of critical thinking on the academic performance of executive MBA students. *Thinking Skills and Creativity*, *31*, 275–283. https://doi.org/10.1016/j.tsc.2019.02.002

Danial, M., Sawal, M., & Nurlaela (2018). Development of Chemistry Instructional Tools and Its Effect on Critical Thinking Skills, Metacognition, and Concept Mastery of Students. *Journal of Physics: Conference Series*, *1028*, 012041. https://doi.org/10.1088/1742-6596/1028/1/012041

Deliligka, A., & Calfoglou, C. (2022). Using poetry to foster Critical Thinking and Metacognition in a Primary School EFL context. *Research Papers in Language Teaching and Learning, 12*(1), 167-187.

Ennis, R. H. (2018). Critical Thinking Across the Curriculum: A Vision. *Topoi*, *37*(1), 165–184. https://doi.org/10.1007/s11245-016-9401-4

Ennis, R. H., Millman, J., & Tomko, T. N. (2005). *Cornell critical thinking tests, level X & level Z. Manual* (5th ed.). Midwest Publications.

Ennis, R. H., & Weir, E. (1985). *The Ennis-Weir Critical Thinking Essay Test.* Midwest Publications

Ericson, J. D. (2022). Mapping the Relationship Between Critical Thinking and Design Thinking. *Journal of the Knowledge Economy*, *13*(1), 406–429. https://doi.org/10.1007/s13132-021-00733-w

Eshmirzaeva, M. A. (2020). Personality-oriented Approach to Education and Modern Pedagogical Technologies. *European Journal of Research and Reflection in Educational Sciences, 8*(8), 25-30.

Evans, C. (2020). *Measuring Student Success Skills: A Review of the Literature on Critical Thinking*. National Center for the Improvement of Educational Assessment.

Facione, P. A. (1990a). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction.* The California Academic Press.

Facione, P. A. (1990b). *The California Critical Thinking Skills Test—College Level Technical Report #1 Experimental Validation and Content Validity.* California Academic Press.

Fatmawati, A., Zubaidah, S., Mahanal, S., & Sutopo. (2019). Critical Thinking, Creative Thinking, and Learning Achievement: How They are Related. *Journal of Physics: Conference Series*, *1417*, 012070. https://doi.org/10.1088/1742-6596/1417/1/012070

Fitriana, L. D., Fuad, Y., & Ekawati, R. (2018). Student's Critical Thinking in Solving Open-Ended Problems Based on Their Personality Type. *Journal of Physics: Conference Series*, *947*, 012007. https://doi.org/10.1088/1742-6596/947/1/012007

Franco, A., Marques Vieira, R., & Tenreiro-Vieira, C. (2018). Educating for critical thinking in university: The criticality of critical thinking in education and everyday life*. Journal for Communication Studies, 11*(2), 131-144.

Frost, P., Donahue, P., Goeben, K., Connor, M., Cheong, H. S., & Schroeder, A. (2019). An examination of the potential lingering effects of smartphone use on cognition. *Applied Cognitive Psychology*, *33*(6), 1055–1067. https://doi.org/10.1002/acp.3546

Ghanizadeh, A. (2017). The interplay between reflective thinking, critical thinking, self-monitoring, and academic achievement in higher education. *Higher Education*, *74*(1), 101–114. https://doi.org/10.1007/s10734-016-0031-y

Gunawardena, M., & Wilson, K. (2021). Scaffolding students' critical thinking: A process not an end game. *Thinking Skills and Creativity*, *41*, 100848. https://doi.org/10.1016/j.tsc.2021.100848

Haber, J. (2020). *Critical Thinking*. MIT Press.

Halpern, D. F. (1997). *Critical thinking across the curriculum: A brief edition of thought and knowledge*. Lawrence Erlbaum Associates.

Halpern, D. F. & Sternberg, R. J. (2019). An Introduction to Critical Thinking: Maybe It Will Change Your Life. In R. J. Sternberg & D. F. Halpern (Eds.), *Critical Thinking in Psychology* (pp. 1-9). Cambridge University Press. https://doi.org/10.1017/9781108684354

Harman, R., & Lee, D. (2009). The role of shame and self-critical thinking in the development and maintenance of current threat in post-traumatic stress disorder. *Clinical Psychology & Psychotherapy*, *17*(1), 13-24. https://doi.org/10.1002/cpp.636

Heidari, K. (2020). Critical thinking and EFL learners' performance on textually-explicit, textually-implicit, and script-based reading items. *Thinking Skills and Creativity*, *37*, 100703. https://doi.org/10.1016/j.tsc.2020.100703

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Ibrahim, I., Akmal, N., & Marwan, M. (2021). Problem-based Learning and Student Critical Thinking to Improve Learning Achievement at Private Universities in Aceh. *AL-ISHLAH: Jurnal Pendidikan*, *13*(2), 1142–1151. https://doi.org/10.35445/alishlah.v13i2.698

Imperio, A., Kleine Staarman, J., & Basso, D. (2020). Relevance of the socio-cultural perspective in the discussion about critical thinking. *Journal of Theories and Research in Education, 15*, 1-19. https://doi.org/10.6092/ISSN.1970-2221/9882

Işiklar, S., & Abali-Öztürk, Y. (2022). The effect of philosophy for children (P4C) curriculum on critical thinking through philosophical inquiry and problem solving skills. *International Journal of Contemporary Educational Research*. *9*(1), 130-142. https://doi.org/10.33200/ijcer.942575

Janssen, E. M., Mainhard, T., Buisman, R. S. M., Verkoeijen, P. P. J. L., Heijltjes, A. E. G., van Peppen, L. M., & van Gog, T. (2019). Training higher education teachers' critical thinking and attitudes towards teaching it. *Contemporary Educational Psychology*, *58*, 310–322. https://doi.org/10.1016/j.cedpsych.2019.03.007

Jebreen, I., & Nabot, A. (2021). An analysis of job advertisements for software engineering employability skills and knowledge: Jordan as a case study. *Asia -Pacific Journal of Science and Technology, 26*(4), APST-26-04-16.

Jin, M., & Ji, C. (2021). The correlation of metacognitive ability, self-directed learning ability and critical thinking in nursing students: A cross-sectional study. *Nursing Open*, *8*(2), 936–945. https://doi.org/10.1002/nop2.702

Kline, R. B. (2015). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.

Kusumoto, Y. (2018). Enhancing critical thinking through active learning. *Language Learning in Higher Education*, *8*(1), 45–63. https://doi.org/10.1515/cercles-2018-0003

Leach, S. M., Immekus, J. C., French, B. F., & Hand, B. (2020). The factorial validity of the Cornell Critical Thinking Tests: A multi-analytic approach. *Thinking Skills and Creativity*, *37*, 100676. https://doi.org/10.1016/j.tsc.2020.100676

Ling, M. K. D., & Loh, S. C. (2020). Relationship of creativity and critical thinking to pattern recognition among Singapore private school students. *The Journal of Educational Research*, *113*(1), 59–76. https://doi.org/10.1080/00220671.2020.1716203

Marthaliakirana, A. D., Suwono, H., Saefi, M., & Gofur, A. (2022). Problem-based learning with metacognitive prompts for enhancing argumentation and critical thinking of secondary school students. *Eurasia Journal of Mathematics, Science and Technology Education*, *18*(9), em2148. https://doi.org/10.29333/ejmste/12304

Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology, 93*(3), 568–592. https://doi.org/10.1037/0021-9010.93.3.568

Merma-Molina, G., Gavilán-Martín, D., Baena-Morales, S., & Urrea-Solano, M. (2022). Critical Thinking and Effective Personality in the Framework of Education for Sustainable Development. *Education Sciences*, *12*(1), 28. https://doi.org/10.3390/educsci12010028

Moreira, H., & Canavarro, M. C. (2018). The Association Between Self-Critical Rumination and Parenting Stress: The Mediating Role of Mindful Parenting. *Journal of Child and Family Studies*, *27*(7), 2265–2275. https://doi.org/10.1007/s10826-018-1072-x

Moroz, M., & Dunkley, D. M. (2019). Self-critical perfectionism, experiential avoidance, and depressive and anxious symptoms over two years: A three-wave longitudinal study. *Behaviour Research and Therapy*, *112*, 18–27. https://doi.org/10.1016/j.brat.2018.11.006

Muhammad Raflee, S. S., & Halim, L. (2021). The Effectiveness of Critical Thinking in Improving Skills in KBAT Problem Solving. *Jurnal Pendidikan Sains Dan Matematik Malaysia*, *11*(1), 60-76. https://doi.org/10.37134/jpsmm.vol11.1.6.2021

Mujanah, S., Ardiana, I. D. K. R., Nugroho, R., Candraningrat, C., Fianto, A. Y. A., & Arif, D. (2022). Critical thinking and creativity of MSMEs in improving business performance during the covid-19 pandemic. *Uncertain Supply Chain Management*, *10*(1), 19–28. https://doi.org/10.5267/j.uscm.2021.10.014

Murtadho, F. (2021). Metacognitive and critical thinking practices in developing EFL students' argumentative writing skills. *Indonesian Journal of Applied Linguistics*, *10*(3). https://doi.org/10.17509/ijal.v10i3.31752

Muthén, B., du Toit, S. H. C., & Spisic, D. (1997). *Robust inference using weighted least squares and quadratic estimating equation in latent variable modeling with categorical and continuous outcomes*. Unpublished manuscript.

Muthén, L. K. (2013, January 28). *Multigroup ESEM* [3:21 pm]. Message posted to http://www.statmodel.com/discussion/messages/11/5581.html?1492019681.

Muthén, L. K., & Muthén, B. O. (2017). *Mplus User's Guide* (8th ed.). Muthén & Muthén.

Muthén, L. K., & Muthén, B. O. (2021). *Mplus (version 8.7)* [Computer software]. Muthén & Muthén.

Ng, S.-Y., Cheung, K., & Cheng, H.-L. (2022). Critical Thinking Cognitive Skills and Their Associated Factors in Chinese Community College Students in Hong Kong. *Sustainability*, *14*(3), 1127. https://doi.org/10.3390/su14031127

Okolie, U. C., Igwe, P. A., Mong, I. K., Nwosu, H. E., Kanu, C., & Ojemuyide, C. C. (2022). Enhancing students' critical thinking skills through engagement with innovative pedagogical practices in Global South. *Higher Education Research & Development*, *41*(4), 1184–1198. https://doi.org/10.1080/07294360.2021.1896482

Renatovna, A. G., & Renatovna, A. S. (2021). Pedagogical and psychological conditions of preparing students for social relations on the basis of the development of critical thinking. *Psychology and Education Journal*, *58*(2), 4889–4902. https://doi.org/10.17762/pae.v58i2.2886

Saleh, S. E. (2019). CRITICAL THINKING AS A 21st CENTURY SKILL: CONCEPTIONS, IMPLEMENTATION AND CHALLENGES IN THE EFL CLASSROOM. *European Journal of Foreign Language Teaching, 4*(1), 1-16. https://doi.org/10.5281/ZENODO.2542838

Sari, D. M. M., & Wardhani, A. K. (2020). Critical thinking as learning and innovation skill in the 21st century. *Journal of English Language and Pedagogy*, *3*(2), 27–34. https://doi.org/10.36597/jelp.v3i2.8778

Saud, W. I. (2020). The Relationship between Critical Thinking and Translation Ability of EFL Undergraduate Students. *International Journal of Social Sciences and Educational Studies*, *7*(3), 19-28. https://doi.org/10.23918/ijsses.v7i3p19

Shahzadi, U., Nimmi, & Khan, I. (2020). Exploring the Relationship between Critical Thinking Skills and Academic Achievement. *Sir Syed Journal of Education & Social Research, 3*(1), 236–242. https://doi.org/10.36902/sjesr-vol3-iss1-2020(236-242)

Sholihah, T. M., & Lastariwati, B. (2020). Problem based learning to increase competence of critical thinking and problem solving. *Journal of Education and Learning (EduLearn)*, *14*(1), 148–154. https://doi.org/10.11591/edulearn.v14i1.13772

Song, Y., Lee, Y., & Lee, J. (2022). Mediating effects of self-directed learning on the relationship between critical thinking and problem-solving in student nurses attending online classes: A cross-sectional descriptive study. *Nurse Education Today, 109*. 105227. https://doi.org/10.1016/j.nedt.2021.105227

Teng, W., Ma, C., Pahlevansharif, S., & Turner, J. J. (2019). Graduate readiness for the employment market of the 4th industrial revolution: The development of soft employability skills. *Education + Training*, *61*(5), 590–604. https://doi.org/10.1108/ET-07-2018-0154

Toker, S., & Akbay, T. (2022). A comparison of recursive and nonrecursive models of attitude towards problem-based learning, disposition to critical thinking, and creative thinking in a computer literacy course for preservice teachers. *Education and Information Technologies*, *27*(5), 6715–6751. https://doi.org/10.1007/s10639-022-10906-y

van der Zanden, P. J. A. C., Denessen, E., Cillessen, A. H. N., & Meijer, P. C. (2019). Patterns of success: First-year student success in multiple domains. *Studies in Higher Education*, *44*(11), 2081–2095. https://doi.org/10.1080/03075079.2018.1493097

Verburgh, A., François, S., Elen, J., & Janssen, R. (2013). The Assessment of Critical Thinking Critically Assessed in Higher Education: A Validation Study of the CCTT and the HCTA. *Education Research International*, *2013*, 1–13. https://doi.org/10.1155/2013/198920

Wakelin, K. E., Perman, G., & Simonds, L. M. (2022). Effectiveness of self-compassion-related interventions for reducing self-criticism: A systematic review and meta-analysis. *Clinical Psychology & Psychotherapy*, *29*(1), 1–25. https://doi.org/10.1002/cpp.2586

Watson, G., & Glaser, E. M. (1980). *Watson-Glaser Critical Thinking Appraisal*. Psychological Corporation.

Watson, G., & Glaser, E. M. (2010). *Watson-Glaser II: Technical manual and user's guide*. Pearson.

Werner, P. H. (1991). The Ennis-Weir Critical Thinking Essay Test: An Instrument for Testing and Teaching. *Journal of Reading, 34*(6), 494-495.

Yuan, R., Liao, W., Wang, Z., Kong, J., & Zhang, Y. (2022). How do English-as-a-foreign-language (EFL) teachers perceive and engage with critical thinking: A systematic review from 2010 to 2020. *Thinking Skills and Creativity, 43*, 101002. https://doi.org/10.1016/j.tsc.2022.101002

Živković, S. (2016). A Model of Critical Thinking as an Important Attribute for Success in the 21st Century. *Procedia - Social and Behavioral Sciences*, *232*, 102–108. https://doi.org/10.1016/j.sbspro.2016.10.034