ASSESSING CRITICAL THINKING INTERVENTIONS IN UNIVERSITY STUDENTS: A LITERATURE REVIEW OF INSTRUMENTS

Barbora Kuvikova – Tomas Sollar

doi: 10.18355/PG.2025.14.2.08

Abstract
The development of critical thinking (CT) is essential for students to participate effectively in a modern society. Still, studies show a lack of clarity of concepts, used methods and measures of intervention. This review focuses on measurement instruments used in critical thinking interventions, evaluating their psychometric properties. A literature search was conducted in three scientific databases - Scopus, Web Of Science, PubMed. The data were systematically retrieved in two phases by two independent researchers following the PRISMA flow diagram. Overall, 13 studies were included in the final analysis. The most frequently used instruments were the California Critical Thinking Skills Test (CCTST) (Facione, 1990) for assessing skills and the California Critical Thinking Disposition Inventory (CTDI) for dispositions. The review also identified a concerning trend of poor psychometric reporting across studies.

Key words
instrument, intervention, critical thinking, university

1 Introduction

Critical thinking

Three primary approaches: philosophical, psychological, and educational, introduced by Sternberg (1986), define critical thinking (CT). Earlier philosophical definitions by Dewey (1933) emphasized the cognitive aspect of critical thinking, portraying it as a skill, a set of skills, a mental procedure, or simply rationality. However, Ennis's work (1962) significantly advanced our understanding by introducing the motivational aspect. The APA Delphi Report (1990) defines it as the disposition to CT, including "the ...motivation of a person to think and his honesty when facing bias."

The last century witnessed a surge in interest in improving critical thinking skills, particularly within nursing education. A recent review by Schoute and Alexander (2023) revealed that most CT intervention programs have been conducted in nursing, pedagogy, and psychology. Strong communication is crucial for nurses to effectively interact with patients, colleagues, and other healthcare professionals (Mohebbi et al., 2022). CT and observation skills are paramount for accurate patient diagnosis and assessment, especially when nurses encounter unexpected situations (Bas et al., 2022). These skills are essential for nurses to navigate complex healthcare environments and provide effective patient care (Carvalho et al., 2020; Cheng et al., 2020). In the field of pedagogy, focus on critical thinking is as important a learning outcome, rather than the definition of a concept, with varying interpretations (Song & Cai, 2024; Nelson et al., 2018). Regardless of the specific approach, critical thinking remains valuable for students in higher education (Taghinezhad & Riasati, 2020). Psychology views CT as encompassing cognitive, attitudinal, and metacognitive components that work together for sound critical thinking performance (Rivas et al., 2022).

Critical thinking (CT) is fundamental for effective participation in modern society, but teaching it is challenging (Ku, 2009). Psychologists worldwide have studied CT

and developed assessment methods. However, the multifaceted nature of critical thinking makes measurement a challenge (Ennis, 2003). As Evans (2020) suggests, the definition of CT determines how it's measured.

How is critical thinking typically measured or assessed?

Critical thinking presents a complex challenge for assessment. It is a multifaceted construct encompassing domain-specific and generalizable elements, integrating both cognitive abilities, exemplified by skills in argument analysis and evaluation, and dispositional factors (Facione, 1990; Evans, 2020), such as intellectual openness and reflectivity (Sosu, 2013). While proficiency in the procedural aspects of critical thinking is essential, the absence of a corresponding disposition to engage these skills in relevant contexts precludes an individual's classification as a critical thinker (Butler, 2024). Two primary assessment instruments are employed when assessing CT, regardless of intervention: standardized tests and performance-based assessments. Examples of standardized tests include the Watson-Glaser Critical Thinking Appraisal Test (Watson & Glaser, 1980) and the California Critical Thinking Skills Test (Facione, 1990), and instruments like the California Critical Thinking Inventory assess critical thinking dispositions (Facione, 1992). Performance-based assessments are particularly well-suited to gather evidence of students' development in applying critical thinking skills and dispositions (Ku, 2009; Schoute & Alexander, 2023).

In defining the concept of critical thinking, the desire to measure both cognitive skills and dispositions necessitates careful consideration of the test format (Lai, 2011). The lack of appropriate assessment tools that can demonstrate the growth of students' critical thinking skills hinders the evaluation of the effectiveness of critical thinking enhancement programs (Ku, 2009). While comprehensive reviews exist on CT in general (Lai, 2011; Butler, 2024), CT interventions (Tiruneh et al., 2014; Puig et al., 2019; Schoute & Alexander, 2023), and methods for capturing CT (Liu et al., 2014; Evans, 2020), there is a notable gap in the literature specifically focused on the measurement of critical thinking within intervention programs. A systematic review by Schoute and Alexander (2023) highlights the growing interest in critical thinking interventions over the past decade, making this review particularly timely and relevant. This review aims to identify and analyze assessment tools for evaluating the impact of critical thinking intervention programs in three areas (nursing, pedagogy, and psychology). The psychometric properties of these tools will be thoroughly assessed. Specific research question is: *What instruments measuring critical thinking are used when evaluating intervention, and what are their psychometric characteristics, and how they capture the specific concept of CT?*

2 Methods

Design

A literature review was conducted according to Grant et al. (2009) publication, which described literature review as providing an examination of recent or current literature, and the PRISMA checklist (Page et al., 2021).

Eligibility criteria and search strategy

The search included all studies published in the last decade meeting predefined criteria: 1. published from January 2014 to March 2024; 2. published as an article; 3. written in English; 4. quantitative research design; 5. university/college students; 6. The used instrument must be mentioned in the Abstract. Exclusion criteria were: 1. other types of publications (reviews, etc.), 2. qualitative or mixed-methods used.

A search was conducted in three scientific databases – Scopus, Web of Science, PubMed – on 27th March 2024. The databases were selected based on their institutional availability. In all three databases, the same key terms were used in

search queries: ((student*) AND (college* OR universit*)) AND ((critical thinking OR critical thinking skill* OR critical thinking disposition* OR analytical thinking OR systematic thinking OR cognitive maturity OR creativity thinking OR judgement* OR decision-making OR decision making OR reasoning OR critical reasoning OR critical thinking process* OR metacognition* OR non-routine thinking OR evaluative thinking OR unsubstantiated belief* OR truth seek*) AND (instrument* OR scale* OR measure* OR tool* OR questionnaire* OR survey*)) AND ((intervention* OR program* OR course)). The search was limited to articles at the title/abstract/keywords level, written in English. A total of 4341 papers were found that matched the stated criteria (1778 from Scopus, 2158 from Web of Science, 405 from PubMed).

Study Selection PRISMA flow diagram

The data were systematically retrieved in two phases by two independent researchers (BK, TS), following the PRISMA flow diagram (Figure 1). In both phases program Rayyan QCRI (http://rayyan.qcri.org) was used. After removal of duplicates (1277), a total of 3064 studies were analyzed using titles, abstracts, and inclusion criteria in the first phase. In the second phase, a total of 18 studies were examined further by reading full-texts (5 of the studies did not provide assessable full-text) until agreement between two independent researchers was achieved. Overall, 13 studies were included in the final analysis.

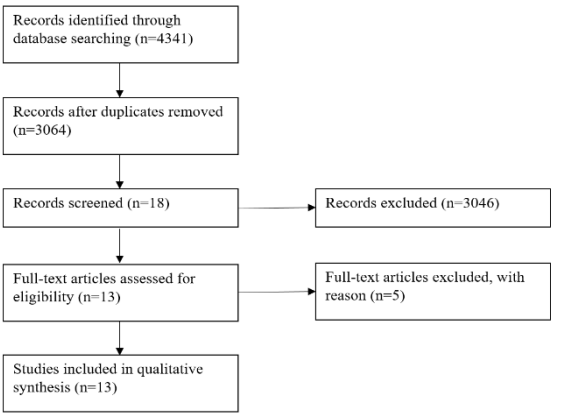

Figure n.1: Flow diagram according to PRISMA

Evaluation of the quality of articles and data extraction

The quality of inclusion was carried out by two independent researchers, using the spreadsheet of their own design (see data extraction). Two independent researchers achieved agreement during the quality evaluation (as recommended in Writing a Psychology Literature Review, 2010). The researchers extracted the following information from the studies: name, author, and year, area of assessment, number of interventions, basic information about intervention, area of research, psychometric information about instrument: number of items, subscales, maximum possible score, type of scale used when answering, reliability, and validity. The synthesis of the data was narrative and tabular.

3 Results

Research question: *What instruments measuring critical thinking are used when evaluating interventions, and what are their basic and psychometric characteristics?*

The literature review identified several instruments used to assess critical thinking (Table 1).

Instrument	Author, year	CT as skills or dispositions	Number of interventions	Intervention		Information about instrument					
				Information about article	Area of research	Nb of items	Subscales	Maximal score or cut of score	Response type	Crombach's alfa	Test-retest
The Critical Thinking Disposition Scale (CTDS)	Semerci (2016)	dispositions	1	The effect of decorative arts course on nursing students' creativity and critical thinking dispositions. Baş, M. T., Ozpulat, F., Molu, B., & Dönmez, H. (2022). Turkey	nursing	49	5: metacognition, well-being, systematicity, tenacity-patience, and openminded ness	Higher mean scores signify a higher critical thinking disposition.	five-point Likert-type scale (1 = Strongly disagree, 2 = Mostly disagree, 3 = Partially agree, 4 = Mostly Agree, 5 = Strongly Agree).	0,96 (provided from different source), 0,91 (from current study)	x
The Critical Thinking Disposition Scale	Ankaraligil (2009)	dispositions	1	The Effect of a Gender Course on the Gender Attitudes, Critical Thinking Dispositions, and Media Literacy Skills of University Students. Koç, G., Özçırpan, Ç. Y., Terzioğlu, F., Çetinkaya, Ş. Ş., Uslu-Şahan, F., Işık, R. A., Başkaya, E. (2021). Turkey	nursing	25	5: researching, self-confidence, flexibility in thought, literacy, and rational thinking.	Higher mean scores signify a higher critical thinking disposition.	five-point Likert-type scale	0,86 (provided from different source), 0,51 (pretest value), 0,75 (post-test value from current study)	X
Ricketts' Critical Thinking Disposition Inventory	Ricketts (2003)	dispositions	1	Synchronous online lecturing or blended flipped classroom with jigsaw: an educational intervention during the Covid-19 pandemic. Mohebbi, Z., Mortezaei-Haftador, A., & Mehrabi, M. (2022). Iran	nursing	33	3: Innovation, perfection and commitment.	x	x	x	x
The California Critical Thinking Disposition Inventory (CTDI)	Facione et al. (1994)	dispositions	2	(1) The effect of blended task-oriented flipped classroom on the core competencies of undergraduate nursing students: a quasi-experimental study, Authors: Ke, L.; Xu, L.; Sun, L.; Xiao, J.; Tao, L.; Luo, Y.; Cao, Q.; Li, Y., 2023, China (2) The Interplay of Critical Thinking Explicit Instruction, Academic Writing Performance, Critical Thinking Ability, and Critical Thinking Dispositions: An Experimental Study. Authors: Taghinezhad, A.; Riasati, M.J.; 2020. Iran	nursing pedagogical	75	7: truth-seeking, open-mindedness, analyticity, systematicity ; self-confidence, inquisitiveness, and cognitive maturity.	Max. 420 points, higher scores indicating higher trends of critical thinking ability	6-point Likert-type scale ranging from 1 (strongly disagree) to 6 (strongly agree)	(1) 0,85 for pre-test and 0,88 for post-test; (2) did not provide this information)	x
The PENCRISAL test	Saiz and Rivas (2008)	skills	1	Metacognitive Strategies and Development of Critical Thinking in Higher Education. Rivas SF, Saiz C and Ossa C (2022). Spain	psychology	35	5: Deductive Reasoning, Inductive Reasoning, Practical Reasoning, Decision-Making, and Problem-Solving	Max score 72	Open-answer format, the person has to answer a concrete question, adding a justification for the reasons behind their answer	0,632 (provided from different source)	0,786 (provided from differed source)
The Critical Thinking Skills Success Technique (Pretest an	Starkey (2010)	skills	1	Interactive learning environment as a source of critical thinking skills for college students. Song, H., & Cai, L. (2024). China	pedagogical	30	x	Max. one point for each question = max. 30	4 options, one is correct	x	x
California Critical Thinking Skills Test	Facione (1990)	skills	3	(1) Critical thinking in nursing students from two Brazilian regions. Carvalho DPSRP, Vitor AF, Cogo ALP, Bittencourt GKGD, Santos VEP, Ferreira Jr MA. (2020). Brazil (2) Specifics of the students' critical thinking formation within active learning space. Solovyeva N, Tapalova O and Smirnov S (2023). Russia* (3) Life Science undergraduate mentors in NE STEM 4U significantly outperform their peers in critical thinking skills. Nelson, K. L., Rauter, C. M., & Cutucache, C. E. (2018). USA	nursing psychology pedagogical	34	8: analysis, interpretation, inference, evaluation, explanation, induction, deduction, and (optional) numeracy	Over 80 points means high level of critical thinking	Multiple choice items use everyday scenarios.	(3) From 0,71 to 0,8 for the subscales; 0,9 overall (provided from different source); studies (1) and (2) did not provide this information	x
The Critical Thinking Task (CTT) test	Miranda (2003)	skills	1	Promoting pedagogy students's scientific skills through critical thinking program. Cornejo, C. O., Luengo, M. P., Mujica, A. D., & Rovira, D. P. (2022). Chile	psychology	8	3: inquiry, analysis and communication; in this study used only dimension "inquiry"	Max. 20 points	Open items	0,75 (provided from different source)	x
Taiwan Critical Thinking Disposition Inventory (TCTDI	Yeh (1998)	critical thinking intention and skills	1	Experiential learning program to strengthen self-reflection and critical thinking in freshmen nursing students during COVID-19: a quasi-experimental study. Cheng, Y. C., Huang, L. C., Yang, C. H., & Chang, H. C. (2020). Taiwan	nursing	20	4: systematicity /analyticity, open-mindedness, inquisitiveness, reflective thinking	'x	6-point Likert-type scale ranging from 1 (least matched) to 6 (highly matched)	From 0,83 to 0,92 (provided from different source), 0,97 (provided from different source)	x

Table n.1: Overview of instruments measuring intervention effect on critical thinking

The CPC2 Critical thinking questionaire	Santiuste-Bermejo et al. (2001)	skills	1	The impact of an educational intervention on nursing students' critical thinking skills: A quasi-experimental study. López, M., Jiménez, J. M., Martín-Gil, B., Fernández-Castro, M., Cao, M. J., Frutos, M., & Castro, M. J. (2020). Spain.	nursing	30	2: substantive dimension of critical thinking in reading, writing, and listening and speaking and dialogic dimension of critical thinking in reading, writing, and listening and speaking	Max. 150 points	1–5 scale being 1: total disagreement, 2: disagreement, 3: sometimes, 4: agreement and 5: total agreement	0,90 (Provided from different source)	x
The Ennis-Weir Critical Thinking Essay Test (CTET)	Ennis and Weir (1985)	skills	1	The Interplay of Critical Thinking Explicit Instruction, Academic Writing Performance, Critical Thinking Ability, and Critical Thinking Dispositions: An Experimental Study. Authors: Taghinezhad, A.; Riasati, M.J.; 2020. Iran	pedagogical	x	general ability to assess an argument and to make an argument in response	x	open-ended test	from 0.72 to 0.99 (provided from different source)	x

Most instruments were published after 2000 (Semerci, 2016; Ankaraligil, 2009; Ricketts, 2003; Saiz & Rivas, 2008; Starkey, 2010; Miranda, 2003; Santiuste-Bermejo et al., 2001). Only four instruments predate 2000 (Facione et al., 1994; Facione, 1990; Yeh, 1998; Ennis & Weir, 1985).

Five instruments focused on measuring critical thinking skills (Saiz & Rivas, 2008; Starkey, 2010; Facione, 1990; Santiuste-Bermejo et al., 2001; Ennis & Weir, 1985). Four instruments focused on dispositions (Semerci, 2016; Ankaraligil, 2009; Ricketts, 2003; Facione et al., 1994). One instrument measured both critical thinking intentions and skills (Yeh, 1998), and one did not specify its area (Miranda, 2003).

Only two instruments in this review were used in more than one intervention: the California Critical Thinking Disposition Inventory (CCTDI) (Facione et al., 1994) (n = 2) and the California Critical Thinking Skills Test (CCTST) (Facione, 1990) (n = 3). Solovyeva, Tapalova, and Smirnov (2023) reported using the CCTDI, but their methodology and results aligned more with the CCTST. Therefore, we classified this study under the California Critical Thinking Skills Test category.

The reviewed studies were conducted primarily in Asia (n = 8). Only two studies originated from Spain (Europe). The remaining studies were conducted in South America (n = 2) and the USA (n = 1). As mentioned in the introduction, the included studies originated from three main areas, with nursing studies being the most common (n = 7).

When considering psychometric information, the number of items in instruments assessing dispositions ranged from 25 to 75. Instruments assessing skills had 20 to 35 items, and one study did not provide this information (Taghinezhad & Riasati, 2020). Also, the number of subscales or dimensions for dispositions varied from three to seven, and for skills, it ranged from two to eight. One study assessing skills did not provide this information (Song & Cai, 2024). Five studies included information about the maximum score achievable. Concerning the evaluation description, three instruments used a five-point Likert scale, while two used a six-point Likert scale. Three instruments used an open-ended format, one used multiple choice, and another used a format with one correct option out of four (for skill assessment).

Reliability was a widely reported psychometric characteristic in our review, with values ranging from 0.51 to 0.99. Two studies (Mohebbi, Mortezaei-Haftador, & Mehrabi, 2022; Song & Cai, 2024) did not provide information on reliability. Three studies mentioned both reliability from the original research and their analysis; one study did not provide this information, and one mentioned another source for reliability and validity evidence. Only one study mentioned evidence for the test-retest reliability of the instrument (Rivas, Saiz, Ossa, 2022).

Assessment tools used across different interventions

The Critical Thinking Disposition Scale (CTDS; Semerci, 2016), a 49-item Likert scale (1-5), assesses five pedagogical dimensions of critical thinking in teachers and candidates. The original scale demonstrated strong reliability (test-retest r = 0.761, Cronbach's $\alpha = 0.963$). In one reviewed intervention (Baş et al., 2022) with nursing

students (n=141, mean age 21.77 ± 0.91 years) aiming to enhance creative thinking with patients, the CTDS reliability was α = 0.91.

Originating from Ankaraligil's (2009) unpublished master's thesis at Istanbul University, information on the 25-item, five-point Likert-type Critical Thinking Disposition Scale (CTDS) comes solely from Koç et al. (2021). The CTDS assesses five critical thinking disposition sub-dimensions, with higher total and subscale means indicating stronger disposition. The original study reported a Cronbach's α of 0.86, lacking validity data. In Koç et al.'s (2021) intervention with nursing students (n=144) on gender role themes, the CTDS Cronbach's α improved in the study group (0.51 to 0.75) and control group (0.64 to 0.77), suggesting some internal consistency improvement.

Inspired by Facione et al.'s (1994) CTDI and originating from Ricketts' (2003) unpublished doctoral thesis, the 33-item Ricketts' Critical Thinking Disposition Inventory assesses three core self-reported dispositions (α = 0.75-0.89; Ricketts & Rudd, 2005). Mohebbi et al. (2022) used an adapted version to study the impact of online learning during COVID-19 on nursing students' critical thinking (n=84, mean age 20.88 years). A separate study by Pakmehr et al. (2013) confirmed the CTDI's validity and reliability in Iran, reporting significant subscale and total score correlations, good item fit in confirmatory factor analysis, and satisfactory split-half reliability.

The PENCRISAL test (Saiz & Rivas, 2008) assesses reasoning, problem-solving, and decision-making in critical thinking, demonstrating acceptable internal consistency (α=0.632) and high test-retest reliability (r = 0.786) in a validation study with Spanish adults (n=715, age 18-53 years). Factor analysis supports its five-factor, 35 open-ended item structure. While showing divergent validity, convergent validity was not established. Filling a gap for Spanish critical thinking assessment focused on processes via real-world scenarios (Rivas & Saiz, 2012), the PENCRISAL was used to measure the effect of a metacognition-focused intervention in Spanish psychology students (n=89, mean age 18.93). Internet administration is possible (Rivas, Saiz, Ossa, 2022).

Originating from Starkey's (2010) book guide, the Critical Thinking Skills Success Technique includes a 30-item, four-option pre-test (one correct answer), 20 critical thinking lessons, and a post-test with an answer key. Its theoretical background assesses specific cognitive skills within a broader critical thinking model. Talov and Orlova's (2020) study found positive correlations with another critical thinking test, suggesting potential validity, but reliability evidence was absent. In our review, Song and Cai (2024) used this technique in an intervention with 60 first-year philology students at Qiqihar University, where participants were randomized into control and Lumosity experimental groups; however, the study did not report validity or reliability data for the Starkey technique.

The California Critical Thinking Skills Test (CCTST) (Facione, 1990), based on Delphi research, assesses core critical thinking skills in college students through 34 multiple-choice items (online/paper-based), measuring 8 dimensions (originally 5). The updated manual (Facione & Facione, 2016) reports established validity (content, construct, criterion) and reliability ($\alpha \geq 0.8$, KR-20: 0.52-0.77). The most used tool in our review (requiring purchase), the CCTST was employed in a Brazilian nursing intervention (Carvalho et al., 2020; n=77) highlighting the potential of high scores for advanced learning and leadership (no reliability/validity data provided). A psychological intervention (Solovyeva, Tapalova & Smirnov, 2023; control n=36, experimental n=36) with psychology and linguistics students, taking 45–50 minutes, lacked clarity on the specific CCTST version used. A US pedagogical study (Nelson et al., 2018) explored undergraduate mentoring's link to critical thinking using the CCTST (no reliability/validity data mentioned).

Developed by Facione et al. (1994) based on the Delphi Report, the California Critical Thinking Disposition Inventory (CTDI) assesses critical thinking dispositions through 70 items across seven subscales (truth-seeking, open-mindedness, analyticity, systematicity, critical thinking self-confidence, inquisitiveness, and cognitive maturity). Respondents use a 6-point Likert scale (total scores 70-420, higher indicating stronger disposition). Requiring purchase and taking approximately 30 minutes (online/paper-pencil administration), the CCTDI exhibits high reliability ($\alpha = 0.90$ overall, subscales 0.71-0.80 in initial study with >1000 students) and strong convergent and divergent validity, with positive and significant inter-dimension correlations. Found across diverse fields and cultures in this review, a nursing study (Ke et al., 2023) using a modified version for Chinese students (n=303) reported good internal reliability ($\alpha = 0.85$-0.88). Conversely, Taghinezhad & Riasati (2020) used the CCTDI without reporting reliability or validity data, limiting generalizability.

The Critical Thinking Task (CTT) test (Miranda, 2003), a Spanish adaptation of the ETS Test (Erwin & Sebrell, 2003), assesses inquiry, analysis, and communication via open-ended items scored 0-6 (max 20 points; Miranda, 2010). While a 2003 study reported good internal reliability ($\alpha = 0.75$) for the inquiry dimension in a Chilean sample (Miranda, 2003), Palma et al. (2021) found lower reliability (0.47-0.59) across dimensions. A Chilean psychological study in our review (Cornejo et al., 2022) focused on inquiry skills using the CTT's 8-item inquiry dimension with pedagogy students (n=95; experimental n=54, control n=41), likely chosen due to prior Chilean application, though their own reliability/validity data was not reported.

A study in Taiwan utilized the Taiwan Critical Thinking Disposition Inventory (TCTDI; Yeh, 1998 in Cheng et al., 2020) to assess critical thinking in nursing students. The original study by Yeh is not publicly accessible online, limiting detailed understanding of its questions and scoring.

The CPC2 Critical thinking questionnaire (Santiuste-Bermejo et al., 2001), originally from an educational practice book, is a 30-item assessment of critical thinking in reading, writing, listening, and speaking. It evaluates reasoning with evidence (substantive) and considering opposing viewpoints (dialogic) on a 1-5 scale (1-2=high difficulty, 4-5=low). The test has high reliability (0.90), though information is limited by accessibility of the original source (Santiuste-Bermejo et al., 2001, in López et al., 2020). Used in a nursing intervention with 112 students, the study (López et al., 2020) advocated for a reflective curriculum to enhance clinical judgment; however, the instrument's reliability was not tested in this intervention.

The Ennis-Weir Critical Thinking Essay Test (CTET; Ennis & Weir, 1985) is a 40-minute, open-ended writing assessment for high school and college students, comprising 9 essays/letters. It evaluates core critical thinking abilities like argument evaluation and viewpoint defense using a structured scoring system (max. 29 points), emphasizing logical reasoning (e.g., sound reasons, assumption identification) over writing mechanics (Ennis & Weir, 1985). The original study reported good content validity and reliability (0.82-0.86). Taghinezhad and Riasati (2020) used the CTET in a pedagogical intervention to assess the impact of critical thinking instruction on medical students' argument analysis and construction, focusing on a philosophical understanding applying formal logic through writing.

4 Discussion

Literature review focused on instruments measuring the effect of interventions on critical thinking conducted in the last 10 years, along with their psychometric characteristics. We identified eleven tools.

These tools primarily focused on critical thinking as skills (Ennis, Weir, 1985; Facione, 1990; Miranda, 2003; Saiz, Rivas, 2008; Santiuste-Bermejo et al., 2001;

Starkey, 2010), while others assessed dispositions to critical thinking (Ankaraligil, 2009; Facione et al., 1994; Ricketts, 2003; Semerci, 2016). Many critical thinking skills assessments overlap in subscales measured (e.g., argument analysis, deductive and inductive reasoning). Conversely, critical thinking disposition tools show wider variation in measured subscales. This distinction is noteworthy for future research and understanding of critical thinking dispositions.

The most frequently used instruments were the California Critical Thinking Skills Test (CCTST) (Facione, 1990) for assessing skills and the California Critical Thinking Disposition Inventory (CTDI) (Facione et al., 1994) for dispositions. Both instruments originated from the APA Delphi research and were used across all three identified research fields: nursing, pedagogy, and psychology.

In last decade, most studies were conducted in Asia (e.g. Ke et al., 2023; Cheng et al., 2020). Additionally, some assessment tools were validated for specific populations, such as Miranda's (2003) Spanish translation and validation. There is a general absence of literature related to investigation of critical thinking interventions in a central or west European context. Our review only identified studies from Spain (López et al., 2020; Saiz, Rivas, Ossa, 2022).

In our review, three of the eleven instruments were initially introduced as part of master's or doctoral theses in pedagogy research (Ankaraligil, 2009; Miranda, 2003; Ricketts, 2003).

Different research areas operationalize critical thinking differently. Facione's (1990) definition, encompassing both critical thinking skills and dispositions, is the most common across fields. While Facione's approach has a philosophical foundation, his research in past was primarily focused on the nursing field. There, critical thinking was associated with creativity and exploring decorative arts (Baş et al., 2022), along with diverse educational approaches like concept map creation (Carvalho et al., 2020; Cheng et al., 2020; Ke et al., 2023; López et al., 2020; Mohebbi et al., 2022) and the role of gender and media literacy (Koç et al., 2021). Psychological studies measured only critical thinking skills, and their background was mostly psychological studies, focused solely on measuring critical thinking skills, drawing primarily from cognitive psychology and differentiating those skills from intelligence. Researchers linked CT skills to metacognition (Rivas, Saiz, Ossa, 2022), probabilistic reasoning skills (Cornejo et al., 2022) and the professional competencies (Solovyeva, Tapalova, Smirnov, 2023 Pedagogical research often provided theoretical introductions, enriching it by current knowledge. The research correlated critical thinking skills and dispositions, as well as academic writing performance of their students (Taghinezhad, Riasati, 2020), others focused on using mobile app when improving critical thinking skills (Song, Cai, 2024) and role of mentoring (Nelson, 2018). Furthermore, according to the previous pandemic, two studies focusing on dispositions of critical thinking were conducted in COVID-19 pandemic context (Mohebbi et al., 2022; Cheng et al., 2020).

Our review also identified a concerning trend of poor psychometric reporting across studies. When searching for data on these properties, we primarily consulted original studies introducing the assessment tools. The most frequently reported psychometric property was reliability, specifically Cronbach's alpha coefficient. While the reviewed studies generally demonstrated acceptable psychometric properties, a gap exists in the literature regarding the inclusion of reliability and validity evidence obtained through the researchers' own investigations (Carvalho et al., 2020; Cornejo et al., 2022; Mohebbi et al., 2022; Nelson et al., 2018; Rivas et al., 2022; Solovyeva et al., 2023; Song & Cai, 2024; Taghinezhad & Riasati, 2020).

Limitations

Review is limited in that it included only studies published in English in databases avaible at the institution where the study was conducted. Only intervention studies

were included in the analysis. Also, many articles were not considered due to not mentioning the assessment tool they used in the abstract of the study.

Acknowledgements
This paper was supported by the University Grant Agency, reg. no. IX/4/2025.

Bibliographic references

Ankaraligil, S. Y. (2009). İlköğretim 6. ve 7. sınıf öğrencilerinde medya okuryazarlığı ve eleştirel düşünme üzerine bir araştırma. Yayımlanmamış Yüksek Lisans Tezi, İstanbul Üniversitesi, İstanbul.

Bas, M. T., Ozpulat, F., Molu, B., & Donmez, H. (2022). The effect of decorative arts course on nursing students' creativity and critical thinking dispositions. Nurse Education Today, 119, 105584. https://doi.org/10.1016/j.nedt.2022.105584

Butler, H. A. (2024). Predicting everyday critical thinking: A review of critical thinking assessments. Journal of Intelligence, 12(2), 16. https://doi.org/10.3390/jintelligence12020016

Carvalho, D. P. D. S. R. P., Vitor, A. F., Cogo, A. L. P., Bittencourt, G. K. G. D., Santos, V. E. P., & Ferreira, M. A. (2020). Critical thinking in nursing students from two Brazilian regions. Revista Brasileira de Enfermagem, 73(1), e20170742. https://doi.org/10.1590/0034-7167-2017-0742

Cornejo, C. O., Luengo, M. P., Mujica, A. D., & Rovira, D. P. (2022). Promoting pedagogy students's scientific skills through critical thinking program. Electronic Journal of Research in Education Psychology, 20 (57), 267-290. https://doi.org/10.25115/ejrep.v20i57.5272

Ennis, R. H. (1962). A concept of critical thinking (p. 84). Harvard educational review.

Evans, C. (2020). Measuring Student Success Skills: A Review of the Literature on Critical Thinking. 21st Century Success Skills. National Center for the Improvement of Educational Assessment.

Facione, P. A. (1990). Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction—The Delphi report. Millbrae, CA: California Academic Press.

Facione, P. A., & Facione, N. C. (1992). The California Critical Thinking Dispositions Inventory. Millbrae, CA: California Academic Press.

Facione, N., & Facione, P. A. (2016). Manual of the California Critical Thinking Skill Test (CCTST).

Facione, N. C., Facione, P. A., & Sanchez, C. A. (1994). Critical thinking disposition as a measure of competent clinical judgement: The development of the California Critical Thinking Disposition Inventory. Journal of Nursing Education, 33(8), 345-350.

Grant, M. J., & Booth, A. (2009). A typology of reviews: an analysis of 14 review types and associated methodologies. Health information & libraries journal, 26(2), 91-108. https://doi.org/10.1111/j.1471-1842.2009.00848.x

Ke, L., Xu, L., Sun, L., Xiao, J., Tao, L., Luo, Y., ... & Li, Y. (2023). The effect of blended task-oriented flipped classroom on the core competencies of undergraduate nursing students: a quasi-experimental study. BMC nursing, 22(1), 1. https://doi.org/10.1186/s12912-022-01080-0

Koc, G., Ozcirpan, C. Y., Terzioglu, F., Cetinkaya, S. S., Uslu-Sahan, F., Isik, R. A., & Baskaya, E. (2021). The effect of a gender course on the gender attitudes, critical thinking dispositions, and media literacy skills of university students. Yükseköğretim Dergisi, 11(2 Pt 2), 387-400. doi:10.2399/yod.20.640377

Ku, K. Y. (2009). Assessing students' critical thinking performance: Urging for measurements using multi-response format. Thinking Skills and Creativity, 4(2009), 70–76.
Lai, E. R. (2011). Critical thinking: A literature review. Pearson's Research Reports, 6(1), 40-41.
Liu, O. L., Frankel, L., & Roohr, K. C. (2014). Assessing critical thinking in higher education: Current state and directions for next-generation assessment. ETS Research Report Series, 2014(1), 1-23. https://doi.org/10.1002/ets2.12009
Lopez, M., Jimenez, J. M., Martin-Gil, B., Fernandez-Castro, M., Cao, M. J., Frutos, M., & Castro, M. J. (2020). The impact of an educational intervention on nursing students' critical thinking skills: A quasi-experimental study. Nurse Education Today, 85, 104305. https://doi.org/10.1016/j.nedt.2019.104305
Miranda, C. (2003). El pensamiento crítico en docentes de educación general básica en Chile: un estudio de impacto. Estudios pedagógicos (Valdivia), (29), 39-54. http://dx.doi.org/10.4067/S0718-07052003000100003
Mohebbi, Z., Mortezaei-Haftador, A., & Mehrabi, M. (2022). Synchronous online lecturing or blended flipped classroom with jigsaw: an educational intervention during the Covid-19 pandemic. BMC Medical Education, 22(1), 845. https://doi.org/10.1186/s12909-022-03915-5
Nelson, K. L., Rauter, C. M., & Cutucache, C. E. (2018). Life Science undergraduate mentors in NE STEM 4U significantly outperform their peers in critical thinking skills. CBE—Life Sciences Education, 17(4), ar54. https://doi.org/10.1187/cbe.18-03-0038
Page, M. J., Mckenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. bmj, 372. doi: https://doi.org/10.1136/bmj.n71

Palma Luengo, M., Ossa Cornejo, C., Ahumada Gutierrez, H., Moreno Osorio, L., & Miranda Jana, C. (2021). Adaptación y validación del test Tareas de Pensamiento Crítico en estudiantes universitarios. Revista de estudios y experiencias en educación, 20(42), 199-212. http://dx.doi.org/10.21703/rexe.20212042palma12
Puig, B., Blanco-Anaya, P., Bargiela, I. M., & Crujeiras-Perez, B. (2019). A systematic review on critical thinking intervention studies in higher education across professional fields. Studies in Higher Education, 44(5), 860–869. https://doi.org/10.1080/03075079.2019.1586333
Ricketts, J. C. (2003). Critical thinking skills of selected youth leaders: The efficacy of leadership development, critical thinking dispositions, and student academic performance (Doctoral dissertation). University of Florida, Gainesville, FL.
Ricketts, J. C., & Rudd, R. D. (2005). Critical thinking skills of selected youth leaders: The efficacy of critical thinking dispositions, leadership, and academic performance. Journal of Agricultural Education, 46(1), 32-43. doi: 10.5032/jae.2005.01032
Fernandez Rivas, S., & Saiz Sanchez, C. (2012). Validación y propiedades psicométricas de la prueba de pensamiento crítico PENCRISAL. Revista electrónica de metodología aplicada, 17(1), 18-34.
Rivas, S. F., Saiz, C., & Ossa, C. (2022). Metacognitive strategies and development of critical thinking in higher education. Frontiers in Psychology, 13, 913219. https://doi.org/10.3389/fpsyg.2022.913219
Saiz, C., And Rivas, S. F. (2008). Evaluación en pensamiento crítico: una propuesta para diferenciar formasde pensar. Ergo. Nueva Época 22-23, 25–66.
Santiuste, V., & Ayala Flores, C. L. (2001). El pensamiento crítico en la práctica educativa. El pensamiento crítico en la práctica educativa.
SEMERCI, N. (2016). The development of critical thinking disposition scale (CTHD): Study on the revision of validity and reliability. International Periodical for

the Languages, Literature and History of Turkish or Turkic, 11(9), 725-740. http://dx.doi.org/10.7827/TurkishStudies.9573
Solovyeva, N., Tapalova, O., & Smirnov, S. (2023, March). Specifics of the students' critical thinking formation within active learning space. In Frontiers in Education (Vol. 8, p. 1132525). https://doi.org/10.3389/feduc.2023.1132525
Song, H., & Cai, L. (2024). Interactive learning environment as a source of critical thinking skills for college students. BMC Medical Education, 24(1), 270. https://doi.org/10.1186/s12909-024-05247-y
Starkey, L. B. (2010). Critical thinking skills success in 20 minutes a day. Learning Express. ISBN: 1576857263. 192 p.
Sternberg, R. J. (1986). Critical thinking: Its nature, measurement, and improvement National Institute of Education. Retrieved from http://eric.ed.gov/PDFS/ED272882.pdf
Schoute, E. C., & Alexander, P. A. (2023). A Critical Analysis of Critical Thinking Interventions in Higher Education. PsychArchives. https://doi.org/10.23668/psycharchives.13935
Taghinezhad, A., & Riasati, M. J. (2020). La interacción de la instrucción explícita del pensamiento crítico, el rendimiento de la escritura académica, la capacidad de pensamiento crítico y las disposiciones de pensamiento crítico: un estudio experimental. IJERI: International Journal of Educational Research and Innovation, (13), 143-165. https://doi.org/10.46661/ijeri.4594
Talov, D. P, Orlova, A. V. (2020). Problems of psychological diagnostics of critical thinking in adolescents. Herzen Readings: Psychol Res Educ.; 3:691–6. doi:10.33910/herzenpsyconf-2020-3-30
Tiruneh, D. T., Verburgh, A., & Elen, J. (2014). Effectiveness of critical thinking instruction in higher education: A systematic review of intervention studies. Higher Education Studies, 4(1), 1-17. doi: 10.5539/hes.v4n1p1
Willingham, D. T. (2007). Critical thinking: Why is it so hard to teach? American Educator, 8–19.
Writing A Psychology Literature Review. (2010). Seattle: University of Washington, Psychology Writing Center. Dostupné z http://www.psych.uw.edu/psych.php#p=339.
Watson, G., & Glaser, E. M. (1980). Watson–Glaser Critical Thinking Appraisal, forms A and B manual. San Antonio, TX: The Psychological Corporation.
Yeh, Y. C. (1998). The basic discussion of effective teaching of critical thinking. Journal of Education Research, 59, 57-67.

Mgr. Barbora Kuviková
Fakulta sociálnych vied a zdravotníctva
Department of Psychological Sciences
Constantine the Philosopher University
Kraskova 1, 94901 Nitra
Slovak Republic
barbora.kuvikova@ukf.sk
ORCID: 0009-0007-9569-6512

prof. PhDr. Tomáš Sollár, PhD.
Department of Psychological Sciences
Constantine the Philosopher University
Kraskova 1, 94901 Nitra
Slovak Republic
tsollar@ukf.sk
ORCID: 0000-0002-8214-896X