

ORIGINAL ARTICLE

Relationships Among Self-Efficacy, Instant Judgement, and Habitual Truth-Digging in Undergraduates' Critical Thinking Disposition

Yong Liu¹  | Attila Pásztor^{2,3}  | Gyöngyvér Molnár^{2,3} 

¹School of Foreign Languages, Chongqing Technology and Business University, Chongqing, China | ²Institute of Education, University of Szeged, Szeged, Hungary | ³MTA-SZTE Digital Learning Technologies Research Group, Szeged, Hungary

Correspondence: Yong Liu (yong.liu@ctbu.edu.cn)

Received: 2 August 2024 | **Revised:** 16 January 2026 | **Accepted:** 21 January 2026

Keywords: critical thinking disposition | Hungarian undergraduate students | validation

ABSTRACT

The Employer-Employee-Supported Critical Thinking Disposition Inventory (2ES-CTDI) bridges theory and practice in higher education by assessing critical thinking disposition (CTD) components: self-efficacy, instant judgement, and habitual truth-digging. However, its broader applicability requires further validation. This study adapted and validated the Hungarian 2ES-CTDI to examine the interactions among these CTD components. A sample of 1090 Hungarian university freshmen completed the 19-item survey (7-point Likert Scale). Reliability and validity were assessed via Cronbach's Alpha, McDonald's Omega, CFA, and Rasch analysis. Results revealed: (a) acceptable reliability (Cronbach's Alpha = 0.784, McDonald Omega = 0.783); (b) overall CFA model fit (CFI and TLI > 0.90, RMSEA and SRMR < 0.08); (c) good item and test endorsements; and (d) relatively low CTD self-efficacy but a strong inclination toward instant judgement. The study provides empirical support for the theoretical CTD framework and offers a validated instrument for cross-cultural comparisons. Practically, it underscores the importance of self-efficacy in fostering higher-order CTD, suggesting its integration into educational standards and curricula. Future research could investigate contextual CTD variations and the broader impacts of self-efficacy interventions. Overall, this work enhances understanding of CTD dynamics and provides insights for educators and policymakers aiming to cultivate critical thinking.

1 | Introduction

The complexity and significance of critical thinking (CT) drive researchers into the continuous exploration of its framework and utilisation, especially in this smart world of the 21st century. People are worried about being replaced by artificial intelligence (AI) machines and chatbots, such as ChatGPT (developed by OpenAI), launched in November 2022, using large language models and reinforcement learning techniques. As a result, it can answer questions, communicate with humans, and create required work. However, according to the category and scope of AI, ChatGPT is still regarded as a weak AI which only contains a limited part of the function of the human mind and can solve narrow tasks (Lock 2022). Besides, most students today are

addicted to AI machines to generate any answers without driving their own thinking, which may lead to the 'thinking laziness' decreasing students' critical thinking and problem-solving competences. Thus, the cultivation and assessment of higher-order thinking skills (HOTS) for students, such as CT, the part of HOTS, are indispensable and vital responsibilities for modern educators and researchers.

Critical thinking disposition (CTD) as the very fundamental elements of CT decides the intrinsic motivation of students to acquire and implement critical thinking skills (CTS) for solving problems (Paul and Binker 1990). However, the research on CTD is in scarcity (Valenzuela et al. 2011), especially the exploration on the newly built CTD framework with the components

of self-efficacy, instant judgement, and habitual truth-digging. This structure and corresponding assessment instrument have been tested in the context of China, while it has not been adopted in Hungarian or other contexts. Thus, this study plans to investigate the feasibility of the instrument and theoretical framework adopted in the Hungarian context, and the deeper relationships among each component.

1.1 | The Framework of CT

CT is the ability to think logically and assess information in a structured way (Alpizar et al. 2022). It is a sophisticated cognitive skill that involves deep mental processes (Ellerton 2022). The goal of critical thinking is to enable people to develop reasonable and thoughtful insights, which in turn helps them to decide what to accept as true or what steps to take. Additionally, it encourages the consideration of suitable options.

The development of the framework of CT can generally be cut into three periods (Liu and Pásztor 2022a; Liu et al. 2026). The CT framework only focused on skills during the first period around the 1970s. It experienced three changes of the focus, from (a) only cognitive skills regarded as problem-solving skills (Kemp 1963), to (b) the combination of judgement and evaluation skills (Quellmalz 1987), and finally to the unity of problem-solving and evaluation skills (Ennis 1987). The second phase, from the 1990s to the 2010s, provided the foundation for CT development and research, broadening the scope of CT to include both cognitive skills and emotionally habitual disposition (Ennis 1996; Facione 1990b; Halpern 1998), which is still used today. Despite lacking the uniqueness of the CT skills (CTS) framework, the widely used one consisting of six dimensions of CTS is considered the source. They are the skills of interpretation, analysis, evaluation, inference, explanation, and self-regulation (Facione 1990a). Other kinds of skills are added in their following studies by researchers, such as creative thinking and mathematical problem-solving skills (Perkins et al. 1993), or verbal reasoning skills (Halpern 1998). Facione and Facione (1992), besides, constructed the other pillar for the framework of CT disposition (CTD), which was summarised from the structure of the Delphi Report (Facione 1990b), considered a relatively stable personality, consisting of seven attributes of inclination: analyticity, inquisitiveness, systematicity, open-mindedness, truth-seeking, self-confidence, and maturity. Based on established theoretical frameworks, Liu and Pásztor (2022a) proposed a comprehensive model that integrates dynamic learning, teaching, and assessment processes. This model is designed to cultivate complex competencies for the AI era, framing critical thinking skills and disposition (CTS/CTD) as the input and real-world problem-solving as the desired output, with targeted instructional interventions connecting theory to practice.

The previous review analysis has proved that instructional intervention is effective for both CTS and CTD (Liu and Pásztor 2022c), which requires educators and researchers to put more attention to their assessments of the growth of CT ability. However, deepening into the CTS and CTD themselves, previous studies demonstrated a high preference for studying the cognitive CTS with the ignorance of CTD and its assessment, which

may be caused by the lack of CTD assessment instruments (Bravo et al. 2020; Quinn et al. 2020; Sosu 2013). However, despite gaining CT capabilities, they are unlikely to be deployed without disposition (Valenzuela et al. 2011). Thus, CTD can be considered the activation of CTS, without which the whole model cannot operate, which drives this study into the investigation of CTD instead of CTS or the overall CT.

1.2 | CTD and Its Assessment

CTD, as the stimulus of CTS and the start point of CT, is a facet of intellectual habit (Paul and Binker 1990) presenting the propensity to think critically. The majority of researchers have concentrated their efforts on examining external elements that affect CTD development in various settings. These include factors like academic success (Liu et al. 2024), self-confidence in academic abilities (Vachova et al. 2023), the state of being mindful (Chen and Tang 2024), the ability to adapt cognitively (Gökçe and Güner 2024), and emotional intelligence (Sk and Halder 2020). While these investigations offer insights into external drivers that can enhance CTD, they have somewhat neglected the intrinsic components of CTD. Delving into the possible interconnections between the different aspects of CTD can assist educators in attaining a more profound comprehension of the CTD structure and the dynamics within its core elements (Zhao et al. 2024).

The framework of CTD, starting with the summarisation of the elements in the Delphi Report (Facione 1990b), develops with the improvement of its assessment instruments. Delphi Report constructed 19 aspects of CTD. They are inquisitiveness, well-informed judgements, open-mindedness, alertness to use CT, prevision, trust in reasoned inquiry, self-confidence, opinions understanding, flexibility in alternatives, fair-mindedness, honesty in facing biases, diligence in judgements, views revision, clarity in statements, working with complexity, reasonableness in criteria, attention on concern, persistence in difficulties, and prudence in judgements. In view of some overlaps, the California Critical Thinking Disposition Inventory (CCTDI; Facione and Facione 1992) narrowed these into seven domains: analyticity, inquisitiveness, systematicity, open-mindedness, truth-seeking, self-confidence, and maturity, which are the reference for following frameworks in inventories.

Foundational instruments like CCTDI have demonstrated considerable cross-cultural structural stability, as evidenced by recent successful validations in contexts such as Hong Kong (Ip et al. 2000). However, parallel research also underscores significant cultural nuances, representing a well-established line of inquiry, consistently affirming the need to examine the contextual robustness of theoretical constructs. This is exemplified by recent adaptations for diverse populations, such as university students in Peru, where rigorous psychometric testing including multigroup analysis confirmed measurement invariance across genders and academic disciplines (Meregildo-Gómez et al. 2025), and for Peruvian adolescents, where the unidimensional structure of a disposition scale was upheld (Martínez-Huamán et al. 2025). These studies, alongside validations in contexts like Hong Kong (Ip et al. 2000), collectively demonstrate that while core CTD constructs show

translatability, their empirical stability and factorial expressions must be ascertained within specific cultural and educational settings. Additionally, Ricketts (2003), on the basis of the preliminary dimensions (willingness of application of prior knowledge and evaluation) of the Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich et al. 1993), built the scheme of innovation, maturity, and engagement. Halpern (2006) set the cognitive-contextual mode contributing to the CTD framework. Yoon (2004) almost renamed the elements of CCTDI into objectivity, prudence, systematicity, intellectual curiosity, intellectual fairness, healthy skepticism, and self-confidence. A suppression into two dimensions of critical openness and reflective skepticism of the Critical Thinking Disposition Scale (CTDS; Sosu 2013) was tested by Bravo et al. (2020) proving that the integration of the two into only one factor seemed better statistically.

Despite the improvements in CTD structures and scales, the previous ones were all based on the theoretical frameworks of classical instruments. Accordingly, a structural map from the perspectives of the previous theories, students, and educators, the Student-Educator Negotiated Critical Thinking Dispositions Scale (SENCTDS) was drawn by Quinn et al. (2020), in which reflection, attentiveness, open-mindedness, organisation, perseverance, and intrinsic goal motivation were included. SENCTDS laid the foundations for the framework construction of the Employer-Employee-Supported Critical Thinking Disposition Inventory (2ES-CTDI; Liu and Pásztor 2022a), with the factors of self-efficacy, instant judgement, and habitual truth-digging, which contained both the intelligence from the workforce and that from previous theories, students, and teachers, to connect the universities with job markets, filling the gap between them that the competence acquired in schools cannot be suitable to work, which is the reason why this study will adopt it as the potential instrument.

1.3 | The Employer-Employee-Supported Critical Thinking Disposition Inventory (2ES-CTDI)

Unlike traditional CTD instruments developed primarily from the perspectives of experts, educators, and students, the 2ES-CTDI was designed to bridge the theory-practice gap. It innovatively integrates insights from the workforce (employers and employees) into its theoretical framework, alongside established structures from prior scales (Liu and Pásztor 2022a). This unique combination aims to create an assessment tool with broader relevance and updated content that reflects the competencies required in real-world professional settings.

The structure of 2ES-CTDI contains the critical-mindedness-orientated disposition consisting of the intrinsic motivational aspects of self-efficacy and the extrinsic perception of instant judgement, and the critical skills-orientated disposition of habitual truth-digging. Liu and Pásztor (2022a) have checked the reliability and validity of 2ES-CTDI in Chinese samples. After the comparisons among one-, two-, and three-factor models, the three-factor models showed the best model fit with self-efficacy, instant judgement, and habitual truth-digging as the framework. It also depicted good reliability (Cronbach's Alpha > 0.70)

and structural validity (convergent and discriminant validity) for both the whole scale and subdimensions.

While a few studies have partially investigated the internal elements of CTD through diverse assessment instruments (Huang et al. 2021; Rauscher and Badenhorst 2020; Sireerat et al. 2022; Chen et al. 2024), they have mainly looked at the disparities in average scores of these elements without delving into the interconnections between them. It was only Liu and Pásztor (2023a) and Zhao et al. (2024) who have undertaken thorough empirical studies within the 2ES-CTDI framework, synthesising and clarifying the uniformities and diversities in the relationships among CTD's components within the Chinese context.

Dispositional self-efficacy represents students' intrinsic motivation: self-confidence, self-regulation, persistence, diligence, and maturity, all of which are linked to individual thoughts about one's ability to govern motivation, cognitive abilities, and actions (Abusalim et al. 2024; Bandura and Jourden 1991; Rayyan et al. 2023), such as to be self-confident and trusting in one's own reasoned judgments; to acknowledge one's sense of self while considering problems and arguments (i.e., life experiences, knowledge, biases, culture, and environment); to be self-efficacious in leading others in the rational resolution of problems; and to recognize that good reasoning is the key to living a rational life and creating (Dwyer et al. 2017). Usher and Pajares (2008) justified policy-level integration of self-efficacy cultivation into national standards, advocating for curriculum reforms prioritizing experiential learning to foster analytical confidence and demonstrating how self-efficacy beliefs mediate students' engagement with complex cognitive tasks. Although it seemed closely correlated with dispositional actions, such as Bandura (1999)'s social cognitive theory, a cornerstone of motivated, intentional learning behaviors, underscored self-efficacy as a lever for enhancing higher-order CTD, some studies (e.g., Quinn et al. 2020) also concluded that self-efficacy was insignificantly correlated with other CTD components. Thus, the role of self-efficacy within CTD was inconsistent.

Extrinsic perception in the instant judgement is represented by using fairness, systematicity, and open-mindedness to analyse the statement quickly and accurately—that is, to make quick judgements on outside statements. This ability to make quick decisions may be handed on to teachers as a safe bet when assessing CT (Leslie and Moilanen 2010). The modern concept of CT described CT as the deliberate and self-regulating judgement that resulted in interpretation, analysis, evaluation, and inference, as well as an explanation of the evidentiary, conceptual, methodological, criteriological, or contextual factors that underpinned that judgement (Facione 1990b), which illustrates that CT, to some extent, should be a slow thinking procedure to collect evidence to make judgements. However, in the human mind, there are both quick and slow thinking processes (Kahneman 2011), and the fast-rhythm and smart society drives people to be efficient. Prior work has overlooked the interplay between self-efficacy and instant judgement tendencies, such as the cognitive biases described by Kahneman (2011), which may impede deliberate, reflective CTD. Thus, the instant judgement should be a part of CTD

with the intelligence accumulation to be a certain subconsciousness from the slow thinking (the consciousness) to promote quick judgement with rationales instead of impulsiveness. This gap is critical, as dual-process theories suggest that intuitive 'System 1' thinking (Evans and Stanovich 2013) often competes with analytical 'System 2' processes central to truth-digging—the relentless pursuit of evidence-based conclusions (Bromme and Goldman 2014).

The emphasis of habitual truth-digging is the inclination to use CTS to seek truth eagerly and ask questions honestly, for example, to think step by step, evaluate with evidence, appraise ideas, and so on (Facione et al. 1997), which is a kind of slowing thinking disposition to collect information domestically and internationally, and systematically and routinely, to put forward opinions evidently, to judge statements accurately, to solve problems effectively, to evaluate consequences thoroughly, and to predict possibilities promisingly. Although this habitual truth-digging was regarded as the main and essential part of CTD, owing to its sluggishness, it should be mastered firstly as a consciousness to have the disposition to think critically, and then with the gradual acquisition into a standard of quantity, this habit needs to be transferred into internal subconsciousness to think and judge quickly and correctly (Liu and Pásztor 2023a). Thus, this balance and transference between instant judgement and habitual truth-digging are a long-lasting endeavour for researchers to explore for the betterment of cultivating undergraduates' CTD; in addition, to master CTS and solve complex scenario problems in daily life (Liu and Pásztor 2022c).

Developing the 2ES-CTDI framework, Liu and Pásztor (2022a) first analysed how the internal elements of this model interacted within a group of 661 Chinese university students. Their research revealed that quick judgement had a positive effect on self-efficacy but a negative impact on the tendency to habitually seek truth. In contrast, self-efficacy was found to markedly promote the development of habitual truth-seeking behaviour. Extending the work of Liu and Pásztor's (2023a) work, Zhao et al. (2024) broadened the study to include 1535 Chinese university students and reconfigured the research model. Their findings challenged those of Liu and Pásztor's (2023a), suggesting that self-efficacy had a substantial negative influence on habitual truth-digging, although quick judgement, acting as a competitive mediator, partially lessened this adverse effect. Both studies by Liu and Pásztor (2023a) and Zhao et al. (2024) also investigated gender disparities in these dynamics, discovering that males might excel in decision-making and judgement tasks, aligning with certain prior research (Reiter 2013; Singh et al. 2022). However, it's crucial to acknowledge that these studies were conducted with Chinese samples, and the relationships among CTD components in other nations or regions are still uncertain.

While the 2ES-CTDI offers an integrated framework for assessing CTD, its development and validation have so far been confined to Chinese contexts. This geographical and cultural limitation presents a critical theoretical gap: it remains unclear whether the proposed tripartite structure (self-efficacy, instant judgement, habitual truth-digging) and the relationships among its components are stable constructs or artefacts of a specific

cultural-educational environment. Therefore, validating the instrument in a distinctly different context, such as Hungary, is not merely an exercise in replication but a necessary test of the framework's theoretical robustness and cross-cultural validity. This study directly addresses this gap by investigating whether the core theoretical propositions of the 2ES-CTDI hold beyond the setting of its origin, thereby contributing to debates on the universality versus situatedness of critical thinking dispositions.

1.4 | CTD Assessment in Hungary

Interest in CT in Hungary was sparked in the late 1990s, following the implementation of the Reading and Writing for Critical Thinking program (RWCT; Klooster et al. 2001). While this program involved creating and applying training materials in classroom contexts (Bárdossy et al. 2002), no sound assessment instruments were developed or applied at the time. A subsequent study developed an assessment tool to measure CT skills; the results among 7th and 11th-grade students revealed poor performance in this domain (Molnár 2002). Further studies have focused on critical reading (Tóth 2007), teacher attitudes toward teaching CT in everyday practice (Fábián 2014; Kaposi 2020), and the assessment of CTS embedded in History education (Érsek 2020). Hungary was also one of 11 countries that participated in the Organisation for Economic Cooperation and Development- Centre for Educational Research and Innovation (OECD-CERI) project aimed at fostering students' creativity and critical thinking (Vincent-Lancrin et al. 2019). While the primary aims of the project were to develop pedagogical resources and carry out interventions, instruments were also developed to measure the effects of the program, such as a scale for learning dispositions related to critical thinking. Results showed that Hungarian students' self-evaluation of CT was one of the highest among the participating countries (Vincent-Lancrin et al. 2019).

This historical overview reveals a persistent and theoretically significant gap in Hungary's approach to CT. Despite early pedagogical initiatives (e.g., RWCT) and participation in international projects (e.g., OECD-CERI), the development and application of robust, domain-general assessment instruments have lagged behind. The existing landscape is fragmented, with studies focusing on specific skills (e.g., critical reading), subject domains (e.g., History), or teacher attitudes, but lacking a cohesive framework for measuring the underlying dispositional components of CT. Notably, the discrepancy highlighted by the OECD-CERI project—where Hungarian students' high self-evaluation of CT contrasted with earlier findings of poor performance—underscores a critical theoretical and methodological issue: the potential disconnect between perceived competence (self-efficacy) and demonstrable skills or dispositions. This gap points directly to the need for a validated instrument that can dissect CT into its core components, such as self-efficacy, judgement, and truth-seeking. Therefore, the present adaptation of the 2ES-CTDI addresses more than a simple lack of tools; it engages with this fundamental debate about the structure and self-perception of CTD. By testing whether this comprehensive framework can be reliably applied in the Hungarian higher education context, this study critically examines if the previously observed disconnect can be traced to specific, measurable facets of the CTD construct.

1.5 | The Research Aims, Questions, and Hypotheses

This study aims to test the reliability, validity, the item and test endorsements of the Hungarian version of 2ES-CTDI (2ES-CTDI-HU) and the relationships among each component for future use in the Hungarian context. Accordingly, on the basis of literature review with the aforementioned research gaps, the research questions (RQs) and hypotheses (Hs) are listed below.

RQ1. What is the feasibility of 2ES-CTDI for Hungarian undergraduates?

H1. *The reliability, validity, and the item and test endorsements of 2ES-CTDI-HU will be acceptable within the Hungarian context. Since the theoretical framework (Liu and Pásztor 2022a), despite a newly built one, was established based on previous theories with the mature accumulation (Bravo et al. 2020; Facione 1990b; Quinn et al. 2020; Sosu 2013), content validity (Liu and Pásztor 2022a), and large-scale assessment in China (Liu and Pásztor 2022a, 2023a, 2023b; Zhao et al. 2024).*

RQ2. What are the relationships among self-efficacy, instant judgement, and habitual truth-digging within the Hungarian context?

H2. *The correlation of these three elements of CTD will be achieved since the theoretical structure was an established one with times investigations on the correlations (Bravo et al. 2020; Facione 1990b; Liu and Pásztor 2022b, 2023a, 2023b; Quinn et al. 2020; Sosu 2013; Zhao et al. 2024). However, the level of each component for Hungarian undergraduates may be different because of the cultural distinctions (Wu 2019; Wu and Molnár 2021) and the scarcity of CTD explorations within the Hungarian context.*

2 | Methods

2.1 | Participants

The study included freshers who had been admitted to a large university in Hungary. The university comprises 12 divisions, such as arts, science, medicine, and others, with varying numbers of students in each division. Out of the total target population of 4068 students, 1090 students (26.8%) took part in the assessment (Mean Age = 19.52, SD = 1.795), of which there were 477 males (Mean Age = 19.61, SD = 1.833) and 613 females (Mean Age = 19.45, SD = 1.763). The participation rate varied across different divisions, ranging from 10.5% to 62.1%.

2.2 | Instrument

2ES-CTDI is a 7-point Likert scale from strongly disagree (1) to strongly agree (7) with 19 items including eight reversed items (A1–A5 and B1–B3), which can be completed by students within 10 min. The 19 items were distributed into three dimensions: self-efficacy (L1, five items), instant judgement (L2, three items), and habitual truth-digging (L3, 11 items). The detailed contents were listed in Table 1. The original version of 2ES-CTDI was in

Chinese. However, for broad and international use, it has been published in the English version (Liu and Pásztor 2022a), and the Chinese version can be requested from the developers. The Hungarian version this study adopted was translated from the English version. The forward-and-back translation approach was carried out for its accuracy. Four experts in English-Hungarian translation participated in this procedure, in which two were set as the forward translation from English to Hungarian, and then, two did the back translation from Hungarian to English. Then, the consistency of the English version after this procedure was checked to ensure the adequacy of the translation so that Hungarian students can understand the contents and expressions well. The Hungarian version can be asked for usage from the correspondent author of this article.

2.3 | Procedures

The CTD assessment was embedded in the competence assessment of the university measuring students' various skills such as reading comprehension, mathematical thinking, or problem-solving (Csapó and Molnár 2017). Students joined the assessment voluntarily and anonymously at the end of August and the beginning of September 2022. They were informed about the opportunity to participate in the assessment. To motivate them, they were given credits upon passing the tests. The online data collection took place in a spacious computer room within the university library via the eDia system (Molnár and Csapó 2019), supervised by test administrators. The study obtained ethical approval from the Institutional Review Board (IRB) of the Doctoral School of Education, University of Szeged (Approval number: 11/2021).

2.4 | Data Analysis

The validation procedure starts with the basic check on factor loadings and the reliability of the scale on each item and factor. The reliability showed in this study adopted the values of Cronbach's Alpha and McDonald's Omega. Because the impact of the number of items on the dependability of Alpha is critical, Omega will be recognised as reliable evidence. Nevertheless, the cutoffs of both are almost the same. The ideal value of them is more than 0.70, but a value above 0.50 can also be considered satisfactory (Cronbach 1951; McNeish 2018; Taber 2018). In addition, the factor loadings need to be above 0.40 (Worthington and Whittaker 2006).

Confirmatory factor analysis (CFA) was applied to test the overall model fit based on the Structural Equation Modelling (SEM) technique. The referred values and cutoffs are: the Chi-square (Pearson 1900) and the Chi-square/df should be below 5; the Root Mean Square Error of Approximation (RMSEA; Steiger 1990) and the Standardised Root Mean Square Residual (SRMR; Bentler 1995) should be less than 0.08 (Hu and Bentler 1999); the Comparative Fit Index (CFI; Bentler 1990) and the Tucker–Lewis index (TLI; Bentler and Bonett 1980; Tucker and Lewis 1973) should be around or more than 0.09. The CFA was designed to test the fit of the single, a priori three-factor model derived from the original 2ES-CTDI theory (Liu and Pásztor 2022d). This model was not compared to alternative

TABLE 1 | The contents of 2ES-CTDI.

Factor	Item	Contents
Self-efficacy (L1)	A1	If, for one thing, most people agree with the conclusion and few reject it, I tend to choose to agree
	A2	When dealing with complicated issues, I often tend to give up
	A3	I'm easily distracted when dedicated to a task
	A4	When making judgements, I am easy to be influenced by ethical values while ignoring reasoning thinking
	A5	When making decisions, I tend to be impulsive
Instant judgement (L2)	B1	The decisions of people with power will be right
	B2	In my mind, critical thinking focuses on criticising others' opinions
	B3	When discussing in a group, if someone's opinion is regarded as wrong by others, he/her should not have any chance to express his/her thoughts
Habitual truth-digging (L3)	C1	If I need to reject others' opinions, I tend to propose my reasons for them
	C2	Whenever I read or hear an assertion or conclusion, I tend to think about possible alternatives
	C3	Before deciding, I tend to search for all the potentially relevant materials
	C4	I tend to consider the social context and reality when dealing with problems
	C5	I prefer proposing suspensions while others are in a calm mood
	C6	I like to analyse any statements or opinions rather than just read and let them go
	C7	I like to deal with problems logically (step by step or with a good plan)
	C8	When judging others' statements, I tend to use reasonable evidence or criteria rather than basing them on my own experience
	C9	I like to reflect on my own or others' thinking procedures
	C10	When I found myself or others using less-persuasive evidence to support opinions, I tend to drive myself or them to adjust the evidence
	C11	I am accustomed to predicting possible results when I solve problems

Note: Items A1–A5 and B1–B3 are reversed items.

structures (e.g., bifactor or higher-order models) as the study's primary objective was the initial cross-validation of this specific theoretical framework, not an exploratory re-specification of its factorial design. Furthermore, chi-square is sensitive to sample size (Babyak and Green 2010) and the estimator this study used was the weighted least square mean and variance adjusted (WLSMV; Muthén 1984). The WLSMV is a reliable estimator that does not require normally distributed variables and is ideal for modelling categorical or ordered data (Brown 2015; Proitsi et al. 2011), and the model fit comparison via chi-square will not work under the estimation of WLSMV. Thus, this study will not consider chi-square and its ratio with df as criteria.

On the basis of the Classic Test Theory (CTT) results, this study also adopted Rasch analysis of the item response theory (IRT; Samejima 1969). The two methods serve distinct but complementary purposes in this validation. The CFA primarily tests the structural validity of the hypothesised three-factor model by examining the relationships between observed items and their posited latent constructs (factors), as well as the correlations among the latent constructs themselves. In contrast, the

Rasch analysis is applied to evaluate measurement quality at the item and scale level. It provides detailed diagnostics on item fit, difficulty, and discrimination, and assesses how well the scale's difficulty targets the ability distribution of the sampled population (i.e., scale targeting). This combined approach offers a more comprehensive understanding of the instrument's psychometric properties, from overall factor structure to individual item performance.

Rasch analysis of the approach of IRT was conducted to assess the test and items endorsement of the inventory. This analysis can generate the infit and outfit mean-square (MNSQ), discrimination, the difficulty of items, and the matching degree between the difficulty of items and students' ability. The infit MNSQ represents the examination of the sensibility of the response patterns in persons, while the outfit MNSQ represents the examination of the sensibility of the response with difficulty far from persons. Although MNSQs depend more on sample size, it is productive for measurement within the range of 0.5 to 1.5 (Linacre 2002). The discriminations with a value of more than 0.4 can be regarded as an acceptable level and those

TABLE 2 | The factor loadings and reliabilities.

Item	L1	L2	L3	Cronbach's α	McDonald's ω
A1	0.504			0.590	0.601
A2	0.862				
A3	0.969				
A4	0.656				
A5	0.618				
B1		0.363		0.506	0.527
B2		0.794			
B3		0.757			
C1			0.641	0.818	0.822
C2			0.825		
C3			0.839		
C4			0.734		
C5			0.264		
C6			0.793		
C7			0.736		
C8			0.565		
C9			0.800		
C10			0.849		
C11			0.680		
Overall reliability				0.784	0.783

Note: L1 (A1–A5): Self-efficacy; L2 (B1–B3): Instant judgement; L3 (C1–C11): Habitual truth-digging; all p -values of items are <0.001 , CI = 95%.

above 0.2 can be the borderline (Ebel 1954). Besides, the difficulty match between items and participants will be shown in the item-person map.

3 | Results

3.1 | Reliabilities and Model Fit

The results of the two kinds of reliabilities and factor loadings are shown in Table 2. Reliability α values of L1, L2, and L3 were 0.590, 0.506, and 0.818, and ω values of them were 0.601, 0.527, and 0.822, respectively. As the influence of the number of items of α on reliability is vital, ω will be regarded as the stable evidence for reliability. Nevertheless, the ω of factor two was still lower than the cutoff of 0.6, indicating the reliability of this dimension was not good. Back to the overall reliabilities, α (0.784) and ω (0.783) were almost the same, revealing the better reliability of the whole inventory.

Besides, factor loadings of all the items in each latent variable were significant with p -values <0.001 (CI = 95%), indicating the items as manifest variables are correlating with the latent constructs that they are measuring (see Table 2). However, some low loadings could also be seen, such as the factor loadings of B1 (0.363) and C5 (0.264), showing the low-level correlation of these

two for their factors, which may decrease the scale reliability. Accordingly, considering all the elements of reliabilities and factor loadings abovementioned, the 2ES-CTDI-HU is reliable with the need for cautious attention to some flawed items.

The model fit of 2ES-CTDI-HU ($N=1090$, $X^2=921.167$, $df=149$) is acceptable with CFI = 0.921 and TLI = 0.909, RMSEA (0.069, $p<0.001$), and SRMR (0.044) <0.08 . The model is shown in Figure 1. The CFA structures can reveal the relationships among each component. The three correlations among L1, L2, and L3 in the model all positively exist, indicating self-efficacy, instant judgement, and habitual truth-digging can be mutually influenced. Additionally, the coefficients of items C5 ('I prefer proposing suspensions while others are in a calm mood.') and B1 ('The decisions of people with power will be right.') seemed to be outliers with low coefficients, testifying, again, to the low factor loadings affecting the reliability and validity of the scale, which arouses the interest to dig deeper for the item and test endorsement.

The findings of acceptable overall reliability and model fit indices partially support Hypothesis 1, confirming that the psychometric properties of the 2ES-CTDI-HU are largely acceptable within the Hungarian context. However, the lower reliability of the second dimension and the low factor loadings of items B1 and C5 indicate areas for caution and potential refinement.

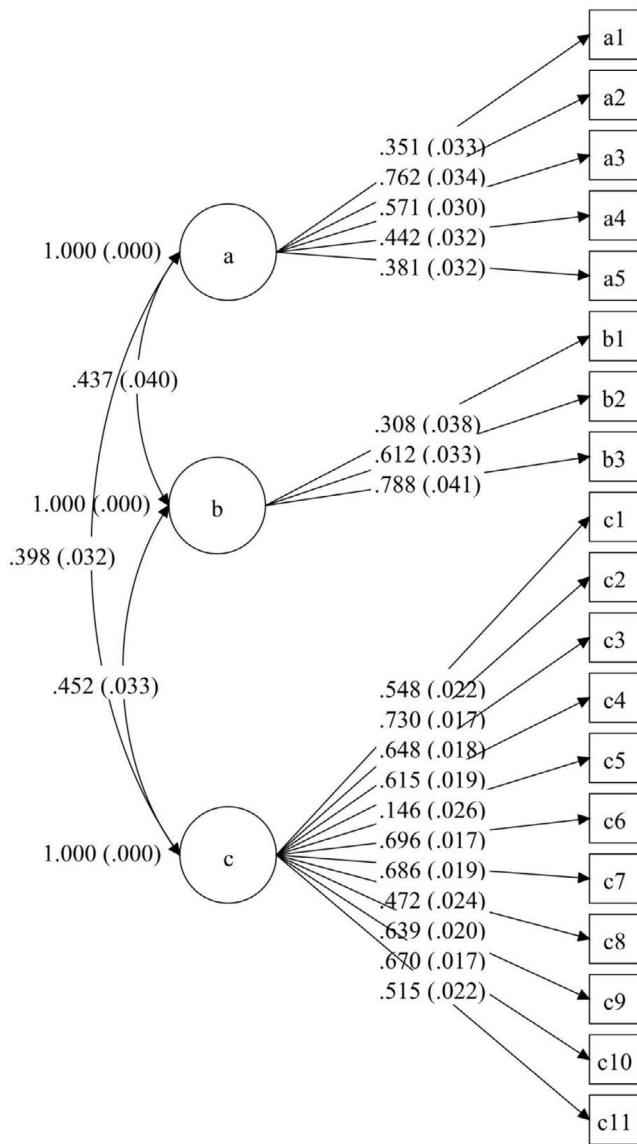


FIGURE 1 | The CFA model of 2ES-CTDI-HU.

3.2 | The Test and Item Endorsement

The test endorsement can be concluded in Figures 2 and 3. The relationship between the student's ability and the expected score showed that the expected score of students' CTD would increase with the growth of students' true ability (theta), indicating the test has a relatively adequate measure of the construct (Figure 2). Figure 3 depicted the relationship between the degree of information provided by the scale and the ability of students, in which the degree of information would increase till the ability of students with $\theta = -2$, and after that, the information that the scale can provide would decrease despite the growth of theta, indicating the scale can provide the most information for students with the CTD ability of -2 level and this scale will be more suitable for the assessment of the lower-average students.

Table 3 showed the infit, outfit, and discrimination of items of 2ES-CTDI-HU. The values of MNSQ infit and outfit of all items were productive for a measurement range of 0.5 to 1.5 within the sample size of this study (Linacre 2002). Most item discriminations are acceptable with a value of more than 0.2 (Ebel 1954),

except item C5, indicating that a good item endorsement existed. Deepen into the item characteristics of C5, Figure 4 demonstrated the possibility of the score students may obtain on item C5, in which the three black vertical lines showed the probability of the attainment of students of scores 2, 4, and 7. For an item with good discrimination, the ability of students should distribute in high dispersion with the possibility of obtaining a score of 1 to 7, and the possibility should increase with the growth of students' competence. However, item C5 cannot accord with this. The possibility of scoring 1 to 7 concentrated within the range around -4 to 0, and with the increase of students' ability, the probability of obtaining higher scores decreased.

To consider items and students' competence together, the difficulty of items and the distribution of persons were shown in Figure 5. The left part of each item-person map represented the distribution of the students (cases), and items were distributed in the right section. The difficulty level decreased from the top to bottom. Therefore, it can be concluded that most of the items are easy for Hungarian undergraduates since a majority of them are located at the bottom, or in other words, most of the students exhibited a tendency to rate themselves highly on the CTD scale. The hardest item was item 5 (A5) and the easiest was item 8 (B3). Together with the item information function distribution in Figure 6, item B3 could provide more information for lower-level students, and item A5 vice versa, also indicating that B3 was the easiest and A5 was the hardest. Besides, almost all the tops of item information curves were distributed at the left part, enhancing the previous conclusion that this scale was suitable for lower CTD level students.

Besides, in Figure 5, most of the harder items belonged to L1 (self-efficacy), such as items 1, 4, 3, and 5 on the top of the item's distribution. Consequently, a conclusion could be made that Hungarian undergraduates rate their CTD self-efficacy lower compared to the other scales. Moreover, the easiest items were 6, 7, and 8 of L2 (instant judgement), which were all listed at the bottom, indicating that they rate their instant judgement ability highly, and the habitual truth-digging items distributed averagely, indicating this disposition was medium and stable.

In summary, the Rasch and item analysis further clarifies the findings for RQ1 and RQ2. Regarding RQ1, the generally productive item fit and good discrimination for most items support the feasibility of the 2ES-CTDI-HU, partially confirming H1, though the poor performance of item C5 indicates a need for revision. For RQ2, the item-person map reveals a distinct profile: items measuring self-efficacy (L1) were among the hardest, while instant judgement (L2) items were the easiest. This pattern strongly supports H2, confirming that the levels of the three CTD components differ within the Hungarian context, with students reporting relatively low self-efficacy but a high inclination toward instant judgement.

4 | Discussions and Conclusion

4.1 | Discussion

The current study advances CTD frameworks by empirically validating the Hungarian 2ES-CTDI, thereby reinforcing the cross-cultural applicability of Facione (1990a) and

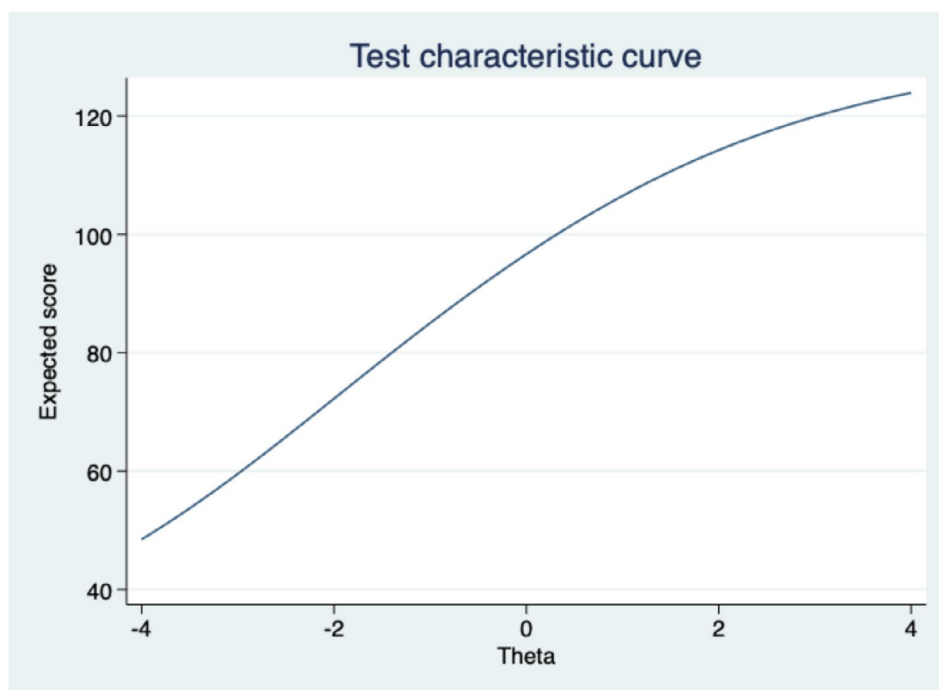


FIGURE 2 | The relationship between the student's ability and the expected score.

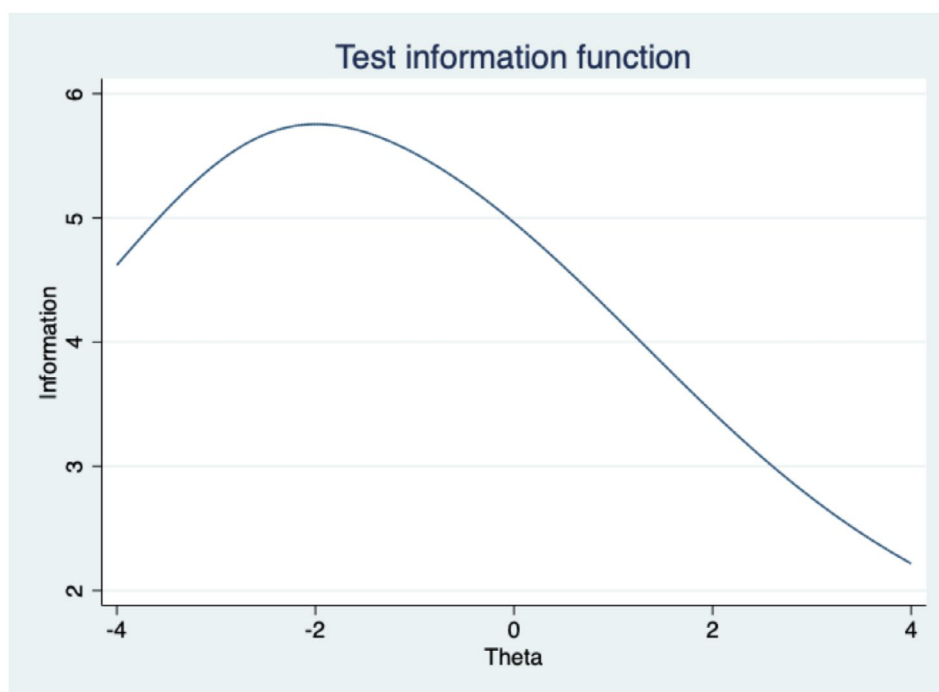


FIGURE 3 | The relationship between the degree of information provided by the scale and the ability of students.

Halpern (2006) theoretical models while addressing critical gaps in the literature. First, our findings challenge assumptions that CTD components operate uniformly across contexts. While the instrument demonstrated structural stability across settings, the specific strength and configuration of relationships among its components (e.g., between self-efficacy and truth-digging) varied. This nuances universalist claims (e.g., Dwyer et al. 2017) and aligns with postcolonial critiques of cognitive frameworks (Kumar 2025) by suggesting that while a core structure of CTD

may be identifiable, its manifestation and developmental pathways are culturally mediated. Consequently, cultivating CTD must account for sociocultural variations in norms related to inquiry and judgement. Second, by confirming Bandura (1999)'s social cognitive theory in the CTD domain, we demonstrate that self-efficacy interventions uniquely enhance analytical rigour—a factor overlooked in prior scale validations but critical for mitigating intuitive 'System 1' biases (Kahneman 2011). This dual theoretical contribution positions the 2ES-CTDI as a

TABLE 3 | The Infit, outfit, and discrimination of items of 2ES-CTDI-HU.

Factor	Item	Number	Discrimination	Infit MNSQ	Outfit MNSQ
Self-efficacy (L1)	A1	1	0.30	1.16	1.17
	A2	2	0.48	0.98	1.02
	A3	3	0.42	1.09	1.11
	A4	4	0.35	1.14	1.17
	A5	5	0.35	1.16	1.17
Instant judgement (L2)	B1	6	0.23	1.14	1.17
	B2	7	0.41	1.04	1.09
	B3	8	0.41	1.01	1.13
Habitual truth-digging (L3)	C1	9	0.47	0.96	0.98
	C2	10	0.63	0.84	0.85
	C3	11	0.59	0.89	0.89
	C4	12	0.55	0.90	0.91
	C5	13	0.18	1.29	1.31
	C6	14	0.64	0.83	0.83
	C7	15	0.63	0.84	0.82
	C8	16	0.47	0.96	0.97
	C9	17	0.56	0.91	0.93
	C10	18	0.60	0.88	0.89
	C11	19	0.48	0.99	1.00

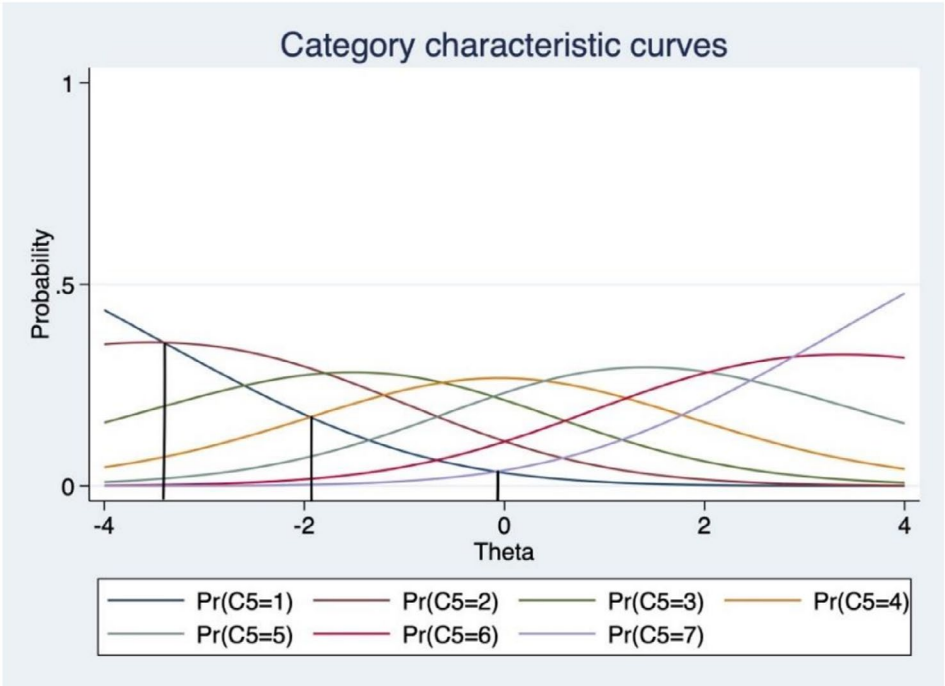


FIGURE 4 | The possibility of the score students may obtain on item C5.

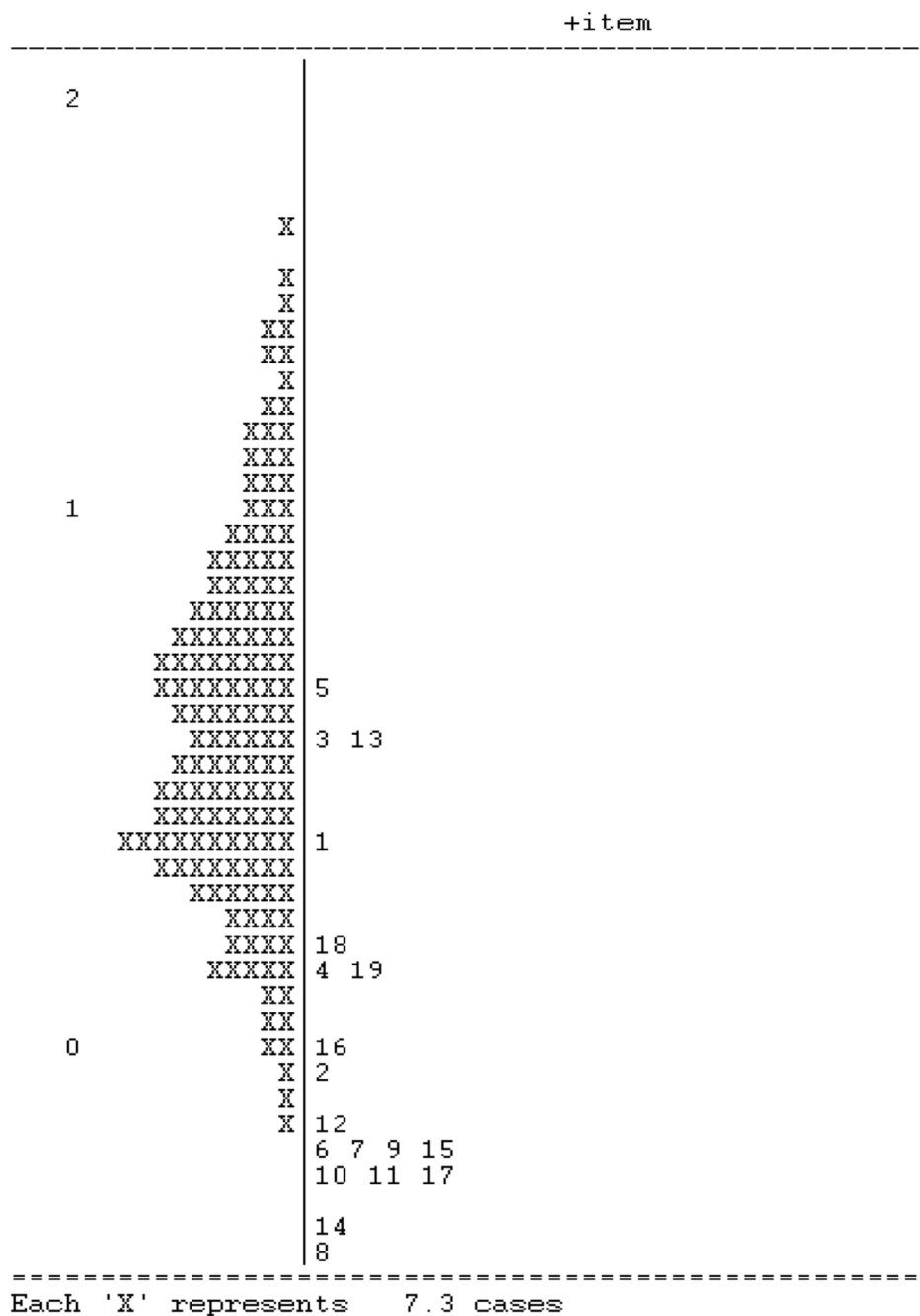


FIGURE 5 | The item-person maps of 2ES-CTDI-HU.

tool for diagnosing and addressing context-specific barriers to reflective thinking.

These findings directly address RQ2 and H2. The confirmation of the tripartite structure supports the existence of correlations among self-efficacy, instant judgment, and habitual truth-digging (aligned with H2). However, the culturally mediated strength and configuration of these relationships, particularly the unique positive link between instant judgment and truth-digging in Hungarians, provide a nuanced answer to RQ2, highlighting significant cultural distinctions in how CTD components interact.

The purpose of this study is to explore the suitability of 2ES-CTDI for Hungarian undergraduate students. The results

showed that the overall reliabilities of both alpha and omega were good enough while L1 and L2 were just acceptable ranging from 0.5 to 0.6. Similar trends can also be found in previous studies even for the most dominant CCTDI in which, for example, the reliabilities of subscales were even around 0.4 or 0.5 (See in Bondy et al. 2001; Ip et al. 2000; Quinn et al. 2020; Walsh and Hardy 1997; Yeh 2002). Deepen into these studies combined with the results of this study, it can be concluded that the low-reliability factors are mostly emotion-orientated elements that probably fluctuate with the change of the current mood of individuals, leading to the low repeatability of subscales. Besides, the low factor loadings of the items in those factors are also the reason causing the inconsistency of the tendency for the responses of items in those factors. For

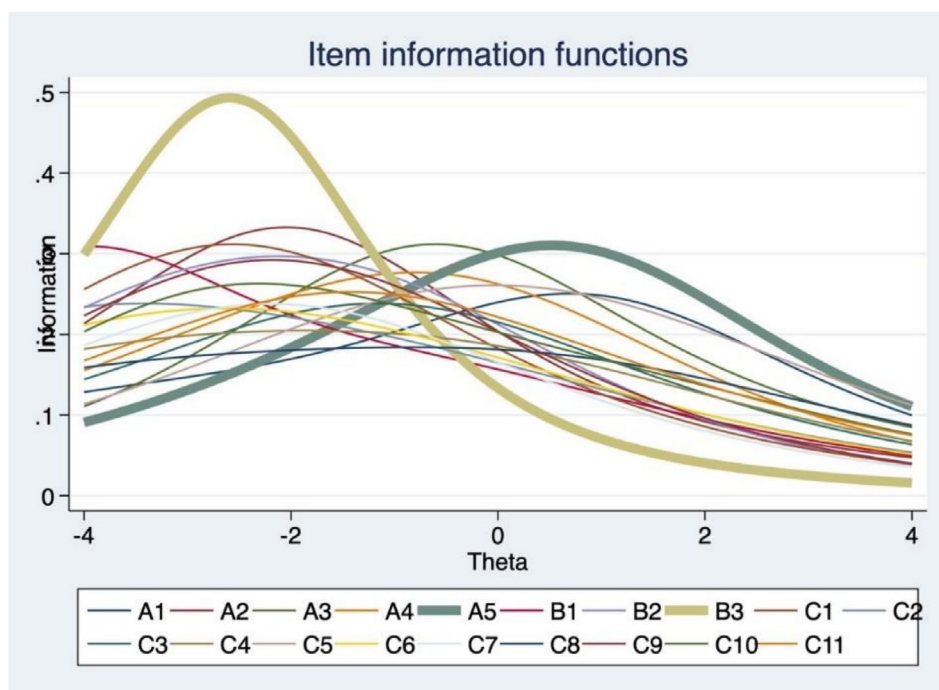


FIGURE 6 | The distribution of information provided by items across students' ability.

instance, in this study, the loadings of items B1 (*The decisions of people with power will be right.*) and C5 (*I prefer proposing suspensions while others are in a calm mood.*) were quite low. Back to their item contents, the wording expressions of them may be cross-loaded with other factors, such as B1 and C5 can both better relate to self-efficacy than the current factor. Thus, most studies put forward the need for CTD scale improvement (see in Kakai 2003; Liu and Pásztor 2022a; Quinn et al. 2020; Wang et al. 2019), but no effective clear measures were suggested.

Collectively, this evidence provides a qualified answer to RQ1 and partial support for H1. While the overall reliability, model fit, and item function are acceptable—supporting the scale's basic feasibility—the suboptimal reliability of certain subscales and the problematic performance of specific items (B1, C5) indicate that H1 is only partially met. This underscores the need for cautious application and targeted refinement of the instrument for the Hungarian context.

The overall model fit, item discriminations, and infit and outfit MNSQs, while acceptable, are at the more modest end of conventional cut-offs. This may be attributed to the complexity of the model, the use of the WLSMV estimator for categorical or ordinal data, and the large sample size which heightens the sensitivity of fit statistics. These results suggest that the three-factor structure is a reasonable but imperfect representation of the data in the Hungarian context. They provide sufficient evidence for the scale's basic structural validity and utility for research, while also indicating that future studies could investigate potential model improvements, such as the strategic parcelling of items from the less reliable 'instant judgement' factor or testing minor residual correlations. Besides, the same tendency happened to items B1 and C5. Although infit and outfit MNSQs were all good, their discriminations were quite low among all the items,

depicting that students' answers tended to be the same for these two items. Nevertheless, the two items can also be included since this scale is to explore the disposition, in other words, it is indeed to find the similar tendency of students. The low discrimination can show that the similarity degree of them is high, and students' opinions are consistent. The difficulty of this scale seemed not quite high, but it is not a test. Moreover, the distribution of the person map was almost normal. As a disposition scale, the difficulty showed most students had a high level of CTD beyond the basic level of this scale, showing Hungarian undergraduate students' rate their CTD high.

The item-person map offers clear empirical support for H2 in relation to RQ2. The systematic difficulty ordering of items—with self-efficacy (L1) being hardest and instant judgement (L2) easiest—provides concrete evidence that the levels of the three CTD components are distinctly different for Hungarian undergraduates, confirming the hypothesised cultural variation in component profiles.

Concerning the results of the correlations, the original validation in Chinese samples demonstrated that there was no correlation between instant judgement and habitual truth-digging (Liu and Pásztor 2022a), while the positive significance of that for Hungarians. For Chinese students, it may be influenced by the Chinese Golden Mean thoughts that students did not make rash decisions and did not hesitate, which demonstrated a balanced inclination to think critically. However, for Hungarians, the teaching of CT has been introduced to some degree in school practice to strengthen students' CT in almost all aspects, including nearly all facets of CTD. Thus, they have both stronger truth-seeking tendencies and the inclination to make fast decisions.

In general, students rate their CTD high. This finding is in line with the findings of the OECD study as well (Vincent-Lancrin

et al. 2019). However, students rate their self-efficacy lower, which may indicate that youngsters prefer to depend on others when making decisions or solving problems.

While this scale validated in Hungarian samples also showed the suitability for lower CTD level students, which is consistent with the results in Chinese samples (Liu and Pásztor 2022a), depicting 2ES-CTDI is more suitable for differentiating lower CTD level students, not for ones with higher proficiency.

The majority of the more difficult items belonged to L1 (self-efficacy), indicating that Hungarian undergraduates rate their CTD self-efficacy lower. Moreover, the simplest item was 8, while items 6, 7, and 8 of L2 (instant judgement) were all placed at the bottom, showing that students had a good inclination for quick judgement. The lack of self-efficacy may indicate that youngsters prefer to depend on others when making decisions or solving problems, while the high tendency to judgement instantly told the truth that they, as adolescents, are impulsive.

In conclusion, this validation study answers both research questions. For RQ1, the 2ES-CTDI-HU is shown to be a feasible, though improvable, tool, partially supporting H1. For RQ2, the distinct profile of component levels and their altered interrelationships confirm that the CTD construct manifests uniquely in Hungary, fully supporting H2 and providing a culturally situated benchmark for future research and intervention.

4.2 | Limitations and Future Directions

While this study provides robust initial evidence for the validity of the 2ES-CTDI-HU, several limitations must be acknowledged to contextualise the findings and guide future research.

First, regarding design and generalisability, the sample consisted exclusively of freshmen from a single Hungarian university. Although the university is large and attracts students from across the country, and the sample size is substantial, the findings may not be fully representative of all Hungarian undergraduates, particularly those in later academic years or from different institutional contexts. The voluntary nature of participation and the varying response rates by academic division introduce the potential for self-selection and non-response bias, which may limit the external validity of the results. Furthermore, the cross-sectional design and reliance on self-report measures restrict the ability to establish causal relationships or track the development of CTD over time. Self-reports may also be susceptible to social desirability biases.

Second, methodological limitations are noted. The model fit indices, while acceptable, were modest, and the reliability of the instant judgement (L2) subscale was below the ideal threshold. This suggests that while the three-factor model is a viable framework, there is room for psychometric refinement. Specific items (e.g., B1, C5) with low factor loadings or discrimination require further scrutiny, as they may reflect cultural nuances in item interpretation, translation issues, or genuine differences in how these constructs manifest in Hungary compared to the original validation context.

Third, the scope of validation evidence is limited. The study did not include tests of measurement invariance. While the model demonstrated acceptable fit in this sample, future research should formally test whether the scale functions equivalently across key groups (e.g., gender) and, crucially, across cultures (e.g., between Hungarian and Chinese students) to justify direct score comparisons. Furthermore, the validation was based primarily on internal structural evidence. The criterion (external) validity of the 2ES-CTDI-HU remains to be established through correlations with objective outcomes such as performance on standardised critical thinking tests, academic grades, or problem-solving tasks in real-world settings.

Consequently, future research should pursue several directions: (1) Multi-institutional and longitudinal studies are needed to enhance the generalisability of the findings and to investigate the developmental trajectories of CTD components. (2) Mixed-methods approaches, combining the scale with interviews or behavioural tasks, could triangulate self-report data and provide richer insights into the cultural and cognitive underpinnings of CTD. (3) Psychometric refinement of the scale, including the potential rewording or replacement of underperforming items and testing of alternative measurement models (e.g., with item parcelling for the L2 factor), would strengthen the instrument for the Hungarian context. (4) Systematic cross-cultural comparisons, building on the preliminary differences observed between Hungarian and Chinese samples, are essential to disentangle the effects of culture, educational systems, and thinking styles on the structure and expression of critical thinking dispositions.

4.3 | Implications

Technically, this study, having demonstrated the complete procedure for validating a questionnaire from the basic loadings, reliability, structural validity, model fit to the item and test endorsement based on both the CTT and IRT instead of the unique CTT adopted in previous studies on scales validation, can be set as a sample for the validation of scales for other studies.

Besides, the robust evidence provided by this study indicated the usability and suitability of the 2ES-CTDI-HU for the assessment of CTD in Hungarian undergraduate students, which will be the starting point for Hungarian universities and educational institutes to conduct the CTD evaluation and fulfil the instructions of the national curriculum.

Theoretically, the three-factor CTD framework model including instant judgement, self-efficacy, and habitual truth-digging has been testified again for its stability, reaffirming the robustness of this practice-informed model. Thus, this theoretical framework, built from the perspectives of multi-aspects, especially the workforce, can be dependable for the assessment of undergraduate students' CTD. However, as item C5 were both with low factor loading and discrimination, the future use needs to treat C5 cautiously, and if necessary, it can be deleted or modified after evaluation. Besides, the instant-judgement (L2) subscale demonstrated modest reliability. Therefore, while the full scale is suitable for group-level research and formative assessment, scores from the L2 subscale should be interpreted with particular caution in high-stakes individual diagnostic contexts. To improve

the instrument, future revisions should consider rewording or replacing item B1, which showed low factor loading and is likely a key contributor to the subscale's reliability issues.

Furthermore, the preliminary findings of the differences in the correlations among each factor for students between Chinese and Hungarian undergraduates can be paid more attention to by educators of both two countries. For example, no correlation between instant judgement and habitual truth-digging for Chinese students while the positive significance of that for Hungarians depicting that the Chinese Golden Mean thoughts that students neither made fast decisions nor hesitated to show a balanced disposition to think critically. However, for Hungarians, the two kinds of tendency were both stronger. Moreover, the lack of self-efficacy requires educators to have deep consideration of it and implement some effective measures to promote students' self-confidence. Nevertheless, the effectively relevant measures need to be explored with the implementation of the instructional actions.

The validated instrument offers actionable pathways for curriculum reform and pedagogical innovation. First, institutions can use the 2ES-CTDI to benchmark CTD competencies across disciplines, enabling targeted interventions. For example, departments with lower truth-digging scores might integrate debate-based assessments or collaborative inquiry projects to foster scepticism (Bromme and Goldman 2014). Second, our findings justify self-efficacy-focused curriculum design: embedding experiential learning modules (e.g., problem-based case studies, simulations) can build confidence in analytical tasks, while meta-cognitive reflection journals help students monitor cognitive biases. Third, the 2ES-CTDI's psychometric rigour allows for longitudinal tracking of CTD development, informing accreditation bodies and policymakers about the efficacy of institutional strategies. For instance, national education standards could mandate CTD assessments alongside traditional metrics, incentivising universities to adopt evidence-based pedagogies. Finally, educators might use the instrument to design bias-awareness training, teaching students to distinguish intuitive judgements from evidence-based conclusions—a skill increasingly vital in misinformation-rich environments.

By bridging theory and practice, this study not only refines CTD frameworks but also equips stakeholders with a diagnostic tool and actionable strategies to cultivate reflective, equitable, and future-ready critical thinkers.

5 | Conclusion

This study validated the Hungarian version of the 2ES-CTDI using a complementary psychometric approach that combined CFA within the CTT framework and Rasch measurement theory. The findings indicate that the adapted instrument is a viable tool for assessing CTD among Hungarian undergraduates. Specifically, the scale demonstrated acceptable overall reliability, and its hypothesised three-factor structure was supported by acceptable model fit indices and significant factor loadings. Rasch analysis further confirmed that the items functioned productively for the target population, revealing that the scale is particularly well-targeted for students with average-to-lower

CTD abilities. A distinctive profile emerged: Hungarian students reported a comparatively lower level of self-efficacy alongside a strong inclination for instant judgement. Based on this evidence, the 2ES-CTDI-HU can be recommended for use in future research and educational assessment to diagnose CTD components within this context. The study also identifies avenues for refinement, such as revising specific items (e.g., B1, C5) with low discrimination to enhance the instrument's precision.

Author Contributions

Yong Liu: conceptualisation; data curation; formal analysis; funding acquisition; investigation; methodology; project administration; resources; software; visualisation; writing – original draft; writing – review and editing. **Attila Pásztor:** supervision; instrument translation administration; data collection; funding acquisition; writing – review and editing. **Gyöngyvér Molnár:** data collection; funding acquisition; writing – review and editing.

Acknowledgements

This is to acknowledge that this work is supported by the Research Start-up Project *The Assessment and Measurement of Undergraduates' Critical Thinking Disposition* for Introduced High-Level Talent of Chongqing Technology and Business University (Grant Number: 2555003), Research Program for Public Education Development of the Hungarian Academy of Sciences: KOZOKT2025-4; Attila Pásztor was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences: BO/00093/23/2; Hungarian National Research, Development and Innovation Fund: OTKAK152413.

Ethics Statement

Participants and the university research committee both gave their consent by ethical standards. Approval Committee: The Institutional Review Board (IRB) of the Doctoral School of Education, University of Szeged. Approval number: 11/2021.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

- Abusalim, N., M. Rayyan, S. Alshanmy, S. Alghazo, and G. Rababah. 2024. "Digital Versus Classroom Discussions: Motivation and Self-Efficacy Outcomes in Speaking Courses via Gather. Town." *Journal of Applied Learning and Teaching* 7, no. 1. <https://doi.org/10.37074/jalt.2024.7.1.24>.
- Alpizar, D., T. Vo, B. F. French, and B. Hand. 2022. "Growth of Critical Thinking Skills in Middle School Immersive Science Learning Environments." *Thinking Skills and Creativity* 46: 101192. <https://doi.org/10.1016/j.tsc.2022.101192>.
- Babak, M. A., and S. B. Green. 2010. "Confirmatory Factor Analysis: An Introduction for Psychosomatic Medicine Researchers." *Psychosomatic Medicine* 72, no. 6: 587–597. <https://doi.org/10.1097/PSY.0b013e3181de3f8a>.
- Bandura, A. 1999. "Social Cognitive Theory: An Agentic Perspective." *Asian Journal of Social Psychology* 2, no. 1: 20. <https://doi.org/10.1111/1467-839X.00024>.

- Bandura, A., and F. J. Jourden. 1991. "Self-Regulatory Mechanisms Governing the Impact of Social Comparison on Complex Decision Making." *Journal of Personality and Social Psychology* 60, no. 6: 941. <https://doi.org/10.1037/0022-3514.60.6.941>.
- Bárdossy, I., M. Dudás, N. C. Pethőné, and R. E. Priskinné. 2002. *A Kritikai Gondolkodás Fejlesztése. Az Interaktív és Reflektív Tanulás Lehetőségei [Developing Critical Thinking. Facilitating Interactive and Reflective Learning]*. University of Pécs.
- Bentler, P. M. 1990. "Comparative Fit Indexes in Structural Models." *Psychological Bulletin* 107, no. 2: 238–246. <https://doi.org/10.1037/0033-2909.107.2.238>.
- Bentler, P. M. 1995. *EQS Structural Equations Program Manual*. Vol. 6. Multivariate software Encino.
- Bentler, P. M., and D. G. Bonett. 1980. "Significance Tests and Goodness of Fit in the Analysis of Covariance Structures." *Psychological Bulletin* 88, no. 3: 588. <https://doi.org/10.1037/0033-2909.88.3.588>.
- Bondy, K. N., L. A. Koenigseder, J. H. Ishee, and B. G. Williams. 2001. "Psychometric Properties of the California Critical Thinking Tests." *Journal of Nursing Measurement* 9, no. 3: 309–328. <https://pubmed.ncbi.nlm.nih.gov/11881271/>.
- Bravo, M. J., L. Galiana, M. F. Rodrigo, J. J. Navarro-Pérez, and A. Oliver. 2020. "An Adaptation of the Critical Thinking Disposition Scale in Spanish Youth." *Thinking Skills and Creativity* 38: 100748. <https://doi.org/10.1016/j.tsc.2020.100748>.
- Bromme, R., and S. R. Goldman. 2014. "The Public's Bounded Understanding of Science." *Educational Psychologist* 49, no. 2: 10. <https://doi.org/10.1080/00461520.2014.921572>.
- Brown, T. A. 2015. *Confirmatory Factor Analysis for Applied Research*. Guilford Publications. <http://www.kharazmi-statistics.ir/Uploads/Public/book/Methodology%20in%20the%20Social%20Sciences.pdf>.
- Chen, A., and Y. Tang. 2024. "Exploring the Facilitating Environment for Mindfulness Meditation: Establishing a Framework through Analysis of Means and Exploratory Factor Analysis." *Mindfulness* 15, no. 10: 2519–2535. <https://doi.org/10.1007/s12671-024-02442-2>.
- Chen, J., X. Wang, and X. Zheng. 2024. "The Investigation of Critical Thinking Disposition Among Chinese Primary and Middle School Students." *Thinking Skills and Creativity* 51: 101444. <https://doi.org/10.1016/j.tsc.2023.101444>.
- Cronbach, L. J. 1951. "Coefficient Alpha and the Internal Structure of Tests." *Psychometrika* 16, no. 3: 297–334. <https://doi.org/10.1007/BF02310555>.
- Csapó, B., and G. Molnár. 2017. "Potential for Assessing Dynamic Problem-Solving at the Beginning of Higher Education Studies." *Frontiers in Psychology* 8: 2022. <https://doi.org/10.3389/fpsyg.2017.02022>.
- Dwyer, C. P., M. J. Hogan, O. M. Harney, and C. Kavanagh. 2017. "Facilitating a Student-Educator Conceptual Model of Dispositions Towards Critical Thinking Through Interactive Management." *Educational Technology Research and Development* 65, no. 1: 47–73. <https://doi.org/10.1007/s11423-016-9460-7>.
- Ebel, R. L. 1954. "Procedures for the Analysis of Classroom Tests." *Educational and Psychological Measurement* 14, no. 2: 352–364. <https://doi.org/10.1177/001316445401400215>.
- Ellerton, P. 2022. "On Critical Thinking and Content Knowledge: A Critique of the Assumptions of Cognitive Load Theory." *Thinking Skills and Creativity* 43: 100975. <https://doi.org/10.1016/j.tsc.2021.100975>.
- Ennis, R. H. 1987. "A Taxonomy of Critical Thinking Dispositions and Abilities." In *Teaching Thinking Skills: Theory and Practice*, edited by J. B. Baron and R. J. Sternberg. W H Freeman/Times Books/Henry Holt & Co.
- Ennis, R. H. 1996. *Critical Thinking*. Prentice-Hall. <https://doi.org/10.5840/inquiryctnews20112613>.
- Érsek, A. 2020. "A Történelmi Tartalomba Ágyazott Kritikai Gondolkodás Kognitív elemeinek Mérési, Fejlesztési Lehetőségei a 11–12 Évfolyamos Gimnáziumi Tanulók Körében Elektronikus Tanulási Környezetben [Assessment and Development Possibilities of the Cognitive Elements of Critical Thinking Embedded in Historical Teaching Materials in the Case of 11th and 12th Graders in an E-Learning Environment]." Dissertation, Eszterházy Károly Catholic University, Eger.
- Evans, J. S. B., and K. E. Stanovich. 2013. "Dual-Process Theories of Higher Cognition: Advancing the Debate." *Perspectives on Psychological Science* 8, no. 3: 223–241. <https://doi.org/10.1177/1745691612460685>.
- Fábián, G. 2014. *Kritikai Gondolkodás az Osztályteremben [Critical Thinking in the Classroom]*. Gondolat Kiadó.
- Facione, P. 1990a. *The California Critical Thinking Skills Test—College Level*. Technical Report# 1. Experimental Validation and Content Validity. <https://files.eric.ed.gov/fulltext/ED327549.pdf>.
- Facione, P. 1990b. "Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction (The Delphi Report)." <https://philarchive.org/archive/FACCTA>.
- Facione, P. A., and N. C. Facione. 1992. *California Critical Thinking Disposition Inventory*. California Academic Press.
- Facione, P. A., N. C. Facione, and C. A. Giancarlo. 1997. *Setting Expectations for Student Learning: New Directions for Higher Education*. California Academic Press.
- Gökçe, S., and P. Güner. 2024. "Pathways from Cognitive Flexibility to Academic Achievement: Mediating Roles of Critical Thinking Disposition and Mathematics Anxiety." *Current Psychology* 43, no. 20: 18192–18206. <https://doi.org/10.1007/s12144-024-05642-0>.
- Halpern, D. 1998. "Teaching Critical Thinking for Transfer Across Domains: Dispositions, Skills, Structure Training, and Metacognitive Monitoring." *American Psychologist* 53: 449–455. <https://doi.org/10.1037/0003-066X.53.4.449>.
- Halpern, D. 2006. *Halpern Critical Thinking Assessment Using Everyday Situations: Background and Scoring Standards*. Unpublished Report.
- Hu, L. T., and P. M. Bentler. 1999. "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives." *Structural Equation Modeling: A Multidisciplinary Journal* 6, no. 1: 1–55. <https://doi.org/10.1080/10705519909540118>.
- Huang, L., A. P.-C. Fan, N. Su, J. Thai, R. O. Kosik, and X. Zhao. 2021. "Chinese Medical Students' Disposition for Critical Thinking: A Mixed Methods Exploration." *BMC Medical Education* 21, no. 1. <https://doi.org/10.1186/s12909-021-02801-w>.
- Ip, W. Y., D. T. F. Lee, I. F. K. Lee, J. P. C. Chau, Y. S. Y. Wootton, and A. M. Chang. 2000. "Disposition Towards Critical Thinking: A Study of Chinese Undergraduate Nursing Students." *Journal of Advanced Nursing* 32, no. 1: 84–90. <https://doi.org/10.1046/j.1365-2648.2000.01417.x>.
- Kahneman, D. 2011. *Thinking, Fast and Slow*. Macmillan. [https://mlsu.ac.in/econtents/2950_Daniel%20Kahneman%20-%20Thinking,%20Fast%20and%20Slow%20\(2013\).pdf](https://mlsu.ac.in/econtents/2950_Daniel%20Kahneman%20-%20Thinking,%20Fast%20and%20Slow%20(2013).pdf).
- Kakai, H. 2003. "Re-Examining the Factor Structure of the California Critical Thinking Disposition Inventory." *Perceptual and Motor Skills* 96, no. 2: 435–438. <https://doi.org/10.2466/pms.2003.96.2.435>.
- Kaposi, J. 2020. "Current Issues Concerning Education for Democracy in Hungary." *Hungarian Educational Research Journal* 10, no. 3: 280–295. <https://doi.org/10.1556/063.2020.00025>.
- Kemp, C. G. 1963. "Improvement of Critical Thinking in Relation to Open-Closed Belief Systems." *Journal of Experimental Education* 31, no. 3: 321–323. <https://doi.org/10.1080/00220973.1963.11010784>.

- Klooster, D. J., J. L. Steele, and P. L. Bloem. 2001. *Ideas Without Boundaries: International Education Reform Through Reading and Writing for Critical Thinking*. International Reading Association.
- Kumar, R. 2025. "Cultivating Anti-Colonicity: A Framework for Building Critical Educational Psychology." *Contemporary Educational Psychology* 102338. <https://doi.org/10.1016/j.cedpsych.2025.102338>.
- Leslie, B. B., and J. Moilanen. 2010. *Advancing Self-Efficacy and Intuitive Decision-Making Calls to Action*. Adult Education Research Conference. <https://newprairiepress.org/aerc/2010/roundtables/8>.
- Linacre, J. 2002. "What Do Infit and Outfit, Mean-Square and Standardized Mean?" *Rasch Measurement Transactions* 16, no. 2: 878. <https://www.rasch.org/rmt/rmt162f.htm#:~:text=Polytomous%20fit%20statistics.,sensitive%20or%20information%2Dweighted%20fit.&text=Outfit%20means%20outlier%2Dsensitive%20fit,person%2C%20and%20vice%2Dversa>.
- Liu, M., T. Vu, N. van Atteveldt, and M. Meeter. 2024. "Correction: Liu et al. (2023). Testing the Reciprocal Effect between Value of Education, Time Investment, and Academic Achievement in a Large Non-Western Sample. *Journal of Intelligence* 11: 1333." *Journal of Intelligence* 12, no. 11: 117. <https://doi.org/10.3390/jintelligence12110117>.
- Liu, Y., and A. Pásztor. 2022a. "Design and Validate the Employer-Employee-Supported Critical Thinking Disposition Inventory (2ES-CTDI) for Undergraduates." *Thinking Skills and Creativity* 46: 101169. <https://doi.org/10.1016/j.tsc.2022.101169>.
- Liu, Y., and A. Pásztor. 2022b. "Design the Employer-Employee-Supported Critical Thinking Disposition Inventory for Undergraduates." In *Sig 1 & 4 Joint Conference 2022, Cadiz, Spain*. https://drive.google.com/file/d/1FebU_u8L30V2D04SueKZ8-KruGkgMz7p/view.
- Liu, Y., and A. Pásztor. 2022c. "Effects of Problem-Based Learning Instructional Intervention on Critical Thinking in Higher Education: A Meta-Analysis." *Thinking Skills and Creativity* 45: 101069. <https://doi.org/10.1016/j.tsc.2022.101069>.
- Liu, Y., and A. Pásztor. 2022d. "Validate the Employer-Employee-Supported Critical Thinking Disposition Inventory for Undergraduates." In *Sig 1 & 4 Joint Conference 2022, Cadiz, Spain*. https://drive.google.com/file/d/1FebU_u8L30V2D04SueKZ8-KruGkgMz7p/view.
- Liu, Y., and A. Pásztor. 2023a. "Moderated Mediating Effects of Gender Among the Components of Critical Thinking Disposition in Undergraduate Students." *Heliyon* 9, no. 4: e14664. <https://doi.org/10.1016/j.heliyon.2023.e14664>.
- Liu, Y., and A. Pásztor. 2023b. "Survey on the Influential Demographic Factors of Chinese Undergraduate Students' Critical Thinking Disposition: Evidence From Plausible Values." *Thinking Skills and Creativity* 50: 101397. <https://doi.org/10.1016/j.tsc.2023.101397>.
- Liu, Y., A. Pásztor, and G. Molnár. 2026. "A Cross-National Assessment on Critical Thinking Disposition Between Hungarian and Chinese Undergraduate Students." *Thinking Skills and Creativity* 60: 102121. <https://doi.org/10.1016/j.tsc.2025.102121>.
- Lock, S. 2022. "What Is AI Chatbot Phenomenon ChatGPT and Could It Replace Humans?" *The Guardian*. <https://www.theguardian.com/technology/2022/dec/05/what-is-ai-chatbot-phenomenon-chatgpt-and-could-it-replace-humans>.
- Martínez-Huamán, H., J. E. Turpo-Chaparro, and C. D. Abanto-Ramírez. 2025. "Psychometric Properties of the Critical Thinking Disposition Scale in Peruvian Adolescents." *Frontiers in Education* 10. <https://doi.org/10.3389/feduc.2025.1537797>.
- McNeish, D. 2018. "Thanks Coefficient Alpha, We'll Take It From Here." *Psychological Methods* 23, no. 3: 412–433. <https://doi.org/10.1037/met0000144>.
- Meregildo-Gómez, R., L.-A. Yglesias Alva, R. Mas-Sandoval, V. Vera-Calmet, H. Aguilar-Armas, and L. Estrada-Yglesias. 2025. "Psychometric Validation and Multigroup Analysis of the Critical Thinking Disposition Scale (CTDS) in University Students." *Journal of Educational and Social Research* 15: 371. <https://doi.org/10.36941/jesr-2025-0220>.
- Molnár, G., and B. Csapó. 2019. "Making the Psychological Dimension of Learning Visible: Using Technology-Based Assessment to Monitor Students' Cognitive Development." *Frontiers in Psychology* 10. <https://doi.org/10.3389/fpsyg.2019.01368>.
- Molnár, L. 2002. "A Kritikai Gondolkodás [Critical Thinking]." In *Az Iskolai Műveltség [Literacy at School]*, edited by B. Csapó, 217–237. Osiris Press.
- Muthén, B. 1984. "A General Structural Equation Model With Dichotomous, Ordered Categorical, and Continuous Latent Variable Indicators." *Psychometrika* 49, no. 1: 115–132. <https://doi.org/10.1007/BF02294210>.
- Paul, R. W., and A. Binker. 1990. *Critical Thinking: What Every Person Needs to Survive in a Rapidly Changing World*. Center for Critical Thinking and Moral Critique, Sonoma State University, Rohnert Park.
- Pearson, K. 1900. "X. On the Criterion That a Given System of Deviations From the Probable in the Case of a Correlated System of Variables Is Such That It Can Be Reasonably Supposed to Have Arisen From Random Sampling." *London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 50, no. 302: 157–175. <https://doi.org/10.1080/14786440009463897>.
- Perkins, D. N., E. Jay, and S. Tishman. 1993. "Beyond Abilities-A Dispositional Theory of Thinking." *Merrill-Palmer Quarterly-Journal of Developmental Psychology* 39, no. 1: 1–21 <Go to ISI>://WOS:A1993KF71500002.
- Pintrich, P. R., D. A. Smith, T. Garcia, and W. J. McKeachie. 1993. "Reliability and Predictive Validity of the Motivated Strategies for Learning Questionnaire (MSLQ)." *Educational and Psychological Measurement* 53, no. 3: 801–813. <https://doi.org/10.1177/0013164493053003024>.
- Proitsi, P., G. Hamilton, M. Tsolaki, et al. 2011. "A Multiple Indicators Multiple Causes (MIMIC) Model of Behavioural and Psychological Symptoms in Dementia (BPSD)." *Neurobiology of Aging* 32, no. 3: 434–442. <https://doi.org/10.1016/j.neurobiolaging.2009.03.005>.
- Quellmalz, E. S. 1987. "Developing Reasoning Skills." In *Teaching Thinking Skills: Theory and Practice*, edited by J. B. Baron and R. J. Sternberg. W H Freeman/Times Books/ Henry Holt & Co.
- Quinn, S., M. Hogan, C. Dwyer, P. Finn, and E. Fogarty. 2020. "Development and Validation of the Student-Educator Negotiated Critical Thinking Dispositions Scale (SENCTDS)." *Thinking Skills and Creativity* 38: 100710. <https://doi.org/10.1016/j.tsc.2020.100710>.
- Rauscher, W., and H. Badenhorst. 2020. "Thinking Critically about Critical Thinking Dispositions in Technology Education." *International Journal of Technology and Design Education* 31, no. 3: 465–488. <https://doi.org/10.1007/s10798-020-09564-3>.
- Rayyan, M., S. Zidouni, N. Abusalim, and S. Alghazo. 2023. "Resilience and Self-Efficacy in a Study Abroad Context: A Case Study." *Cogent Education* 10, no. 1: 2199631. <https://doi.org/10.1080/2331186X.2023.2199631>.
- Reiter, K. K. 2013. *Gender Differences in Decision Making When Faced with Multiple Options*. In Saint John's University. <https://doi.org/10.1016/j.tsc.2023.101444>.
- Ricketts, J. C. 2003. *The Efficacy of Leadership Development, Critical Thinking Dispositions, and Student Academic Performance on the Critical Thinking Skills of Selected Youth Leaders*. University of Florida. <https://ufdc.ufl.edu/UFE0000777/00001/pdf/>.
- Samejima, F. 1969. "Estimation of Latent Ability Using a Response Pattern of Graded Scores." *Psychometrika Monograph Supplement* 34, no. 4: 100.
- Singh, V., S. Thakral, K. Singh, and R. Garg. 2022. "Examining Cognitive Sex Differences in Elite Math Intensive Education: Preliminary

- Evidence from a Gender Inequitable Country.” *Trends in Neuroscience and Education* 26: 100172. <https://doi.org/10.1016/j.tine.2022.100172>.
- Sireerat, K., N. Seki, M. Akiyama, A. Kinoshita, and I. Morio. 2022. “Critical Thinking Disposition Among Thai dental students.” *Journal of Dental Education* 86, no. 8: 968–975. <https://doi.org/10.1002/jdd.12913>.
- Sk, S., and S. Halder. 2020. “Critical Thinking Disposition of Undergraduate Students in Relation to Emotional Intelligence: Gender as a Moderator.” *Heliyon* 6, no. 11: e05477. <https://doi.org/10.1016/j.heliyon.2020.e05477>.
- Sosu, E. M. 2013. “The Development and Psychometric Validation of a Critical Thinking Disposition Scale.” *Thinking Skills and Creativity* 9: 107–119. <https://doi.org/10.1016/j.tsc.2012.09.002>.
- Steiger, J. H. 1990. “Structural Model Evaluation and Modification: An Interval Estimation Approach.” *Multivariate Behavioral Research* 25, no. 2: 173–180. https://doi.org/10.1207/s15327906mbr2502_4.
- Taber, K. S. 2018. “The Use of Cronbach’s Alpha When Developing and Reporting Research Instruments in Science Education.” *Research in Science Education* 48, no. 6: 1273–1296. <https://doi.org/10.1007/s11165-016-9602-2>.
- Tóth, L. 2007. *Kritikai Olvasás—Kritikai Gondolkodás [Critical Reading—Critical Thinking]*. Pedellus.
- Tucker, L. R., and C. Lewis. 1973. “A Reliability Coefficient for Maximum Likelihood Factor Analysis.” *Psychometrika* 38, no. 1: 1–10.
- Usher, E. L., and F. Pajares. 2008. “Sources of Self-Efficacy in School: Critical Review of the Literature and Future Directions.” *Review of Educational Research* 78, no. 4: 45. <https://doi.org/10.3102/0034654308321456>.
- Vachova, L., E. Sedlakova, and J. Kvintova. 2023. “Academic Self-Efficacy as a Precondition for Critical Thinking in University Students.” *Pegem Journal of Education and Instruction* 13, no. 2: 328–334. <https://doi.org/10.47750/pegegog.13.02.36>.
- Valenzuela, J., A. M. Carracedo, and C. Saiz. 2011. “Critical Thinking Motivational Scale: A Contribution to the Study of Relationship Between Critical Thinking and Motivation.” *Electronic Journal of Research in Educational Psychology* 9: 823–848. <https://doi.org/10.25115/ejrep.v9i24.1475>.
- Vincent-Lancrin, S., C. González-Sancho, M. Bouckaert, et al. 2019. “Fostering Students’ Creativity and Critical Thinking.” *Educational Research and Innovation*. <https://doi.org/10.1787/62212c37-en>.
- Walsh, C. M., and R. C. Hardy. 1997. “Factor Structure Stability of the California Critical Thinking Disposition Inventory Across Sex and Various Students.” *Majors* 85, no. 3_suppl: 1211–1228. <https://doi.org/10.2466/pms.1997.85.3f.1211>.
- Wang, X., X. Sun, T. Huang, R. He, W. Hao, and L. Zhang. 2019. “Development and Validation of the Critical Thinking Disposition Inventory for Chinese Medical College Students (CTDI-M).” *BMC Medical Education* 19, no. 1: 200. <https://doi.org/10.1186/s12909-019-1593-z>.
- Worthington, R. L., and T. A. Whittaker. 2006. “Scale Development Research: A Content Analysis and Recommendations for Best Practices.” *Counseling Psychologist* 34, no. 6: 806–838. <https://doi.org/10.1177/0011000006288127>.
- Wu, H. 2019. “Identifying the Influencing Factors of Problem Solving: A Cross-National Study.” Doctoral Dissertation, University of Szeged. <https://doktori.bibl.u-szeged.hu/id/eprint/10333>.
- Wu, H., and G. Molnár. 2021. “Logfile Analyses of Successful and Unsuccessful Strategy Use in Complex Problem-Solving: A Cross-National Comparison Study.” *European Journal of Psychology of Education* 36, no. 4: 1009–1032. <https://doi.org/10.1007/s10212-020-00516-y>.
- Yeh, M. L. 2002. “Assessing the Reliability and Validity of the Chinese Version of the California Critical Thinking Disposition Inventory.” *International Journal of Nursing Studies* 39, no. 2: 123–132. [https://doi.org/10.1016/s0020-7489\(01\)00019-0](https://doi.org/10.1016/s0020-7489(01)00019-0).
- Yoon, J. 2004. *Development of an Instrument for the Measurement of Critical Thinking Disposition in Nursing*. Catholic University of Education.
- Zhao, Y., Y. Liu, and H. Wu. 2024. “Relationships Among Critical Thinking Disposition Components of Chinese Undergraduates: A Moderated Mediating Effect Analysis.” *International Journal of Educational Research* 124: 102306. <https://doi.org/10.1016/j.ijer.2023.102306>.