Heliyon 9 (2023) e14664

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Moderated mediating effects of gender among the components of critical thinking disposition in undergraduate students

Yong Liu [a,*], Attila Pásztor [b,c,d]

[a] Doctoral School of Education, University of Szeged, Hungary
[b] Institute of Education, University of Szeged, Hungary
[c] MTA-SZTE Research Group on the Development of Competencies, Hungary
[d] MTA-SZTE Digital Learning Technologies Research Group, Hungary

ARTICLE INFO

Keywords:
Moderated mediation
Critical thinking disposition
Gender difference
Undergraduate students

ABSTRACT

Critical thinking disposition (CTD), as a facet of the critical thinking (CT) framework, is regarded as the activation of critical thinking skills (CTS). Although studies regarding gender difference of CTD can be found, there is a scarcity of studies on the relationships among CTD components and their mediating effects concerning gender. Additionally, traditional gender comparison of latent means neglected the influence of scales, which leads to doubt about the results on whether the difference is caused by scales or gender characteristic distinction. Measurement invariance (MI) is suggested to be confirmed before performing comparisons. Previous research has had fewer MI results for CTD inventories. Thus, this study aims to analyze the gender equality of the Employer–Employee–Supported Critical Thinking Disposition Inventory (2ES-CTDI) as well as the moderated mediating effects of gender on the components of CTD among 661 Chinese undergraduates (M_{age} = 19.57; SD = 1.26) via MI and Partial Least Squares Structural Equation Modeling (PLS-SEM) by Mplus and STATA. The results found that (a) the scale has high reliability and validity for measuring undergraduates' CTD. MI findings indicated that the configural and metric models were achieved, and the scalar model identified the partial invariance by freeing the intercepts of indicators A5, C7, and C8; (b) Girls have stronger self-efficacy and habitual truth-digging disposition, whereas boys have higher instant judgment; (c) Instant judgment has a negative influence on habitual truth-digging with self-efficacy as the competitive partial mediator, in which, gender moderated the relationship between instant judgment and self-efficacy. These findings, theoretically, proved the stability of the CTD framework of 2ES-CTDI, and practically, call for instructors to pay closer attention to the gender role in CTD cultivation.

1. Introduction

As a component of the critical thinking (CT) framework, critical thinking disposition (CTD) is defined as the activation of critical thinking skills (CTS) [1–3]. Without the foundation of CTD, CTS cannot be mastered well. However, most studies preferred doing research on CTS to CTD because of its lack of assessment instruments. Furthermore, although some studies on CTD can be found, there are few studies on the relationships between CTD components and their gender-mediated effects. The exploration of the gender role is vital, which is an asset in all fields including education. Instructors should know if the gender difference exists in CTD instead of just the

* Corresponding author.
E-mail address: yong.liu@edu.u-szeged.hu (Y. Liu).

https://doi.org/10.1016/j.heliyon.2023.e14664
Received 24 September 2022; Received in revised form 6 March 2023; Accepted 14 March 2023
Available online 20 March 2023

physiological difference to promote the CTD effectively. In addition, traditional group comparisons of latent means ignored the influence of scales, casting doubt on whether the difference is caused by scales or gender characteristic distinction. Thus, before performing comparisons, measurement invariance (MI) should be confirmed. Previous studies found fewer MI results for CTD inventories. Therefore, this study plans to determine the gender role in the relationships between CTD components in undergraduate students on the basis of MI results via a newly developed CTD tool.

2. Literature review

2.1. Framework of critical thinking disposition

CTD, the activation of the cognitive CTS [1–3], is the psychometric section of the CT framework. CTD can be regarded as the tendency, inclination, and willingness to think critically. Even while its framework varies throughout investigations, it largely maintains stability because practically all the CTD frameworks were drawn from the Delphi report [4] and the California Critical Thinking Disposition Inventory (CCTDI) [5].

Though there were some overlaps among them, the broader CTD framework was identified in the Delphi report with 19 factors—namely, inquisitiveness, well-informed, open-mindedness, alternatives to use CT, precision, trust in reasoned inquiry, self-confidence, understanding opinions, flexibility in alternatives, fair-mindedness, honesty in facing biases, diligence in judgments, revise views, clarity in statements, working with complexity, reasonableness in criteria, attention on concern, persistence in difficulties, and prudence in judgments. As a result, the CCTDI, the first instrument for CTD evaluation, focused on seven dimensions that serve as the foundation for all subsequent frameworks: analyticity, inquisitiveness, systematicity, open-mindedness, truth-digging, self-confidence, and maturity.

Pintrich, Smith [6] built the Motivated Strategies for Learning Questionnaire that contained the subscale of CT, the intensive framework of which consisted of applying previous knowledge to new situations and giving evaluations. In his CTD questionnaire, Ricketts [7] integrated the framework into innovation, maturity, and engagement. Yoon [8] made a few adjustments to the framework of CCTDI, setting the framework into objectivity, prudence, systematicity, intellectual curiosity, intellectual fairness, healthy skepticism, and self-confidence. Sosu [9] suppressed the framework into critical openness and reflective skepticism, establishing the critical thinking disposition scale (CTDS), and the adapted version of Bravo, Galiana [10] proved better with the combination of Sosu's two domains into one factor.

The aforementioned frameworks were all influenced by earlier ideas, and two new frameworks were established that were not only based on literature but also input from students and teachers and employers and employees, respectively. Quinn, Hogan [11] tested the framework in the Student–Educator Negotiated Critical Thinking Dispositions Scale, including reflection, attentiveness, open-mindedness, organization, perseverance, and intrinsic goal motivation. Liu and Pásztor [12] set up the framework of self-efficacy, instant judgment, and habitual truth-digging in their Employer–Employee–Supported Critical Thinking Disposition Inventory (2ES-CTDI), which is the newest scale. The dimensions in this framework were obtained from the perspectives of employers and employees on the basis of previous frameworks with the expectation to know what kind of graduates the workforce needs on CTD. The disposition of self-efficacy shows students' intrinsic motivation: self-confidence, self-regulation, persistence, diligence, and maturity, which are related to individual beliefs on the ability to manage motivation, cognitive skills, and actions [13]. Using fairness, systematicity, and open-mindedness to judge the statement quickly and accurately, —that is, to make quick decisions—represents the extrinsic perception in the instant judgment. This ability to make quick decisions may be passed on to educators as a prudent trust when assessing critical thinking [14]. Habitual truth-digging focuses on the inclination to use CTS to seek truth eagerly and ask questions objectively, for example, to think step by step, analyze with evidence, evaluate ideas, and so on [15].

In conclusion, the CTD framework showed variation, but the foundation of them is distinct. They all got inspiration from CCTDI. Therefore, even though the overall CTD results were obtained under different frameworks from different scales, they can still be taken into account collectively. However, since they have different subdimensions within the overall scale, the relationships among each subdimension will be completely distinct. Therefore, given the significance of the exploration of these relationships, this study will carry out the analysis within the newest scale 2ES-CTDI regarding the interactions among self-efficacy, instant judgment, and habitual truth-digging.

2.2. Gender difference in CTD among undergraduates

Most studies adopted CCTDI as the instrument to assess CTD for undergraduates. Walsh and Hardy [16] used CCTDI to summarize that females performed better in open-mindedness and maturity of CTD than males. Yenice [17] found that there was no difference in the overall CTD, but females scored higher in open-mindedness, whereas Elkoca [18] got the reverse result on open-mindedness in nursing students. According to Biber, Tuna [19], there was no gender difference in the overall CTD, although girls outperformed for the subdimension of analyticity. Partial similarity existed in Bilen, Ercan [20]'s study that there was a difference in overall CTD and for the subdimension of analyticity, females performed better. Du, Emmersen [21] and Mwalongo [22] obtained the same tendency for open-mindedness and inquisitiveness. In contrast, males are more systematic and inquisitive, and there was no difference in their open-mindedness, according to the opposing conclusions [23]. The findings of Fitriani, Asy'ari [24] in prospective teachers demonstrated mixed opinions that females scored higher in inquisitiveness and maturity, whereas males performed better in self-confidence and open-mindedness. However, there was no difference between Whitney, Aleksejuniene [25]'s study on dental students and Ding [26]'s study on students studying English as a foreign language in any of the subdimensions.

In addition to these subdimensions of CCTDI, Bers, McGowan [27], Kirbaslar and Ozsoy-Guneş [28], and Li [29] found the overall CTD of females was still higher than males. However, Najafi, Motlagh [30] obtained the opposite results for medical students. Dennett and DeDonno [31] compared college students from America and China and found that American girls tended to be more open-minded than boys, although this was not the case for Chinese students. However, Koyunlu Unlu and Dokme [32], Demirbag, Unisen [33], Özyurt [34], Kilic, Yazici [35], and the meta-analysis of Liu and Pásztor [36] found no difference across gender on CTD.

Some studies used other scales for CTD assessment. Shida, Osman [37] and Mousazadeh, Momennasab [38] adopted the Ricketts' critical thinking disposition questionnaire [7] on polytechnic and nursing students respectively, and the results showed no statistical difference among gender in overall CTD. However, females scored higher significantly in the subdimension of engagement [38], and no difference in the subdimensions [37]. Park [39] used Yoon's critical thinking disposition scale [8] on nursing students, finding the reverse result that males performed better in CTD.

Stupnisky, Renaud [40] adopted the Motivated Strategies for Learning Questionnaire [6] as the CTD scale to compare gender differences under the comparison between the baseline model and path coefficients constrained model without a detailed MI test, concluding that there was no significant difference across gender. With only a one-factor model, Bravo, Galiana [10] conducted the MI across gender on the adapted Critical Thinking Disposition Scale [9], which showed that girls have greater CTD than boys but that the one-factor model was not the best. Álvarez-Huerta, Muela [41] used the adapted Critical Thinking Disposition Scale [10] to determine the gender difference in CTD without the investigation of MI on this scale generating the same conclusion that girls' CTD is higher than boys. However, Liu and Pásztor [42] used 2ES-CTDI without MI showing that no difference was found across gender.

To sum up, just a few studies performed MI before comparing latent means, indicating the absence of MI testing in this sector, whereas most of the earlier studies investigated the gender difference in CTD without taking the MI into account, yielding a variety of results. Additionally, although there have been instances of inconsistencies, most studies have found that females performed better on both the sub-dimensional and overall CTD.

2.3. Mediating and moderating effects among CTD components in undergraduates

For the mediating effects on CTD, for example, Dökmecioğlu, Tas [43] concluded that metacognitive self-regulation strategies mediated the relationship between constructivist learning environment perceptions and CTD. Kim, Seo [44] found that disaster recognition has a significant relationship with CTD with the major satisfaction as the mediator. Cheng and Wan [45] concluded that the classroom learning environment had a stronger relationship with CTD than CTS with CTD as the mediator in this relationship. Song, Lee [46] generated that self-directed learning mediated the relationship between CTD and problem-solving ability in nursing undergraduate students. Álvarez-Huerta, Muela [41] found that openness to diversity and challenge mediated the positive relationship between CTD and undergraduate students' creative confidence beliefs.

For the moderating effects of gender on CTD, for example, Orhan [47] found that gender moderated one's decision-making process based on 48 alpha values from 39 studies in a meta-analysis. Sk and Halder [48] found that emotional intelligence has a positive effect on CTD without the moderation of gender. Ali and Awan [49] and Fong, Kim [50] concluded that gender moderated the relationships between CTD and academic achievement. While Abiogu, Ede [51] and Liu and Pásztor [36] had negative results that gender was not the moderator of the relationships between cognitive-behavioral reflective training or problem-based learning and CTD respectively.

In conclusion, almost no studies explored the moderated mediating effects among the component of CTD with gender as the moderator. On the mediating and moderating effects among CTD components in undergraduate students, few studies have been conducted. Most earlier studies concentrated on how teaching strategies, psychological variables, or other abilities affected CTS and the overall CTD. In addition, previous studies demonstrated that emotional factors can affect human behavior so for a behavioral disposition [52]; the directions of these influences are different, positive or negative. Stedman and Andenoro [53] found significant associations between emotional intelligence skills and CTD in undergraduate leadership students. Thus, this study assumed that the emotional factor (self-efficacy) could be set as the mediator among the components of CTD.

3. The current study

In light of the literature review, the current study intends to identify the gender role in the whole and subdimensions of CTD via the instrument 2ES-CTDI based on MI. Accordingly, considering the necessity of the check of MI before gender difference and influence analyses, this study aims to investigate the reliabilities and model fits on the overall scale and across gender, determine the gender equality of the scale that was employed, explore the variation in latent CTD means across gender based on MI, and figure out the moderated mediating effects of gender across the CTD components. The research questions are as follows:

Research question 1. Are the overall reliabilities and model fit, and those across gender acceptable?

Research question 2. Can this study establish the full scalar invariance or strict invariance model across gender?

Research question 3. Can this study identify the latent difference across gender based on the MI model?

Research question 4. Does gender moderate the relationships among the components of CTD?

4. Method

4.1. Research design

The study design consists of two phases: data collection and data analysis. The data collection was carried out in October 2021 in southwestern China. Four universities agreed to participate in this project. The selected scale was sent to the correspondent student counselors who are in charge of delivering the questionnaire to students via email, social media applications, or on-site in the classroom by scanning the QR code. The total distribution was around 800 undergraduate students, and 661 were submitted back. The rate of recovery was around 83%. The collected data then were retrieved from the online system for analysis.

Based on the research purpose regarding the gender role among the components of CTD, the data analysis procedure started from the exploration of the gender equality of the adopted scale via the MI method. While the invariance was obtained, the latent difference was investigated as the foundation of the moderating role of gender in the mediating effect model. Then, the mediating effect model was established and the influence of the moderator of gender was checked on each path.

4.2. Data collection

4.2.1. Participants

For the recruitment of students, student counselors at four universities in southwestern China were contacted to distribute online questionnaires. Participants and the university research committee both gave their consent by ethical standards (Approval committee: The Institutional Review Board of the Doctoral School of Education, University of Szeged. Approval number: 11/2021). Finally, 661 undergraduate students voluntarily and anonymously participated in the assessment (M_{age} = 19.57; SD = 1.26). The detailed demographic information is listed below in Table 1. Specifically, there were 328 boys (49.6%; M_{age} = 19.22; SD = 1.21) and 333 girls (50.4%; M_{age} = 19.91; SD = 1.24). Based on their majors, two categories were set including 274 undergraduates in Humanity, Arts, and Social Sciences, while 387 in Natural Sciences. Among them, 159 are vocational college students of all three grades, and 502 from all four grades of universities.

4.2.2. Instrument

For evaluating CTD in undergraduate students, the 2ES-CTDI (see Appendix) has been found to have strong reliability, acceptable content, and convergent and discriminant validity [12]. It is a 7-point Likert scale (from strongly disagree to strongly agree) consisting of three dimensions: self-efficacy (five items), instant judgment (three items), and habitual truth-digging (11 items), which were adapted from CCTDI, CTDS, SENCTDS, and the scale developers. It is an online questionnaire, including 19 items in total, that can be delivered via link or QR code, and the duration for filling out this scale can be about five to 8 min. Besides, at the end of the questionnaire, some demographic information questions are listed, such as gender, grade, and major. The collected data will be recorded automatically in the account and can be downloaded in the format of Excel or SPSS for further analysis.

4.3. Data analysis

4.3.1. Measurement invariance and latent mean difference

MI, the premise of the latent mean comparisons [54,55], is an approach to ensure if there is a significant difference among groups, contexts, and time series in the relationships between their latent factors and indicators based on a scale [56]. MI test has a specific order requirement: configural (baseline), metric (weak), scalar (strong), and strict invariance model [57]. However, as shown in Fig. 1, a single-group CFA test should be conducted before MI to investigate the model fit in each group.

Maximum likelihood robust (MLR), as opposed to maximum likelihood (ML), which is utilized for data with a normal distribution, was the estimator used in this investigation. However, in most cases in the real study, the data cannot achieve normality, and ML is not suitable for ordinally scaled data [58]. Thus, MLR [59] with robust corrections to standard errors and chi-square goodness of fit

Table 1
The demographic information of the sample.

Demographic information	Sub-category	N	Percentage (%)	Age Mean ± SD
Gender	Boy	328	49.6	19.22 ± 1.21
	Girl	333	50.4	19.91 ± 1.24
Major	Humanity, Arts, & Social Sciences	274	41.5	20.02 ± 1.26
	Natural Sciences	387	58.5	19.24 ± 1.16
Grade[a]	Vocational 1	108	16.3	18.77 ± 1.04
	Vocational 2	21	3.2	19.76 ± 0.83
	Vocational 3	30	4.5	20.53 ± 1.80
	Bachelor 1	149	22.5	18.66 ± 0.91
	Bachelor 2	196	29.7	19.66 ± 0.82
	Bachelor 3	138	20.9	20.58 ± 0.87
	Bachelor 4	19	2.9	21.14 ± 1.91

Note: a. Vocational level and bachelor level both belong to the university level in China.

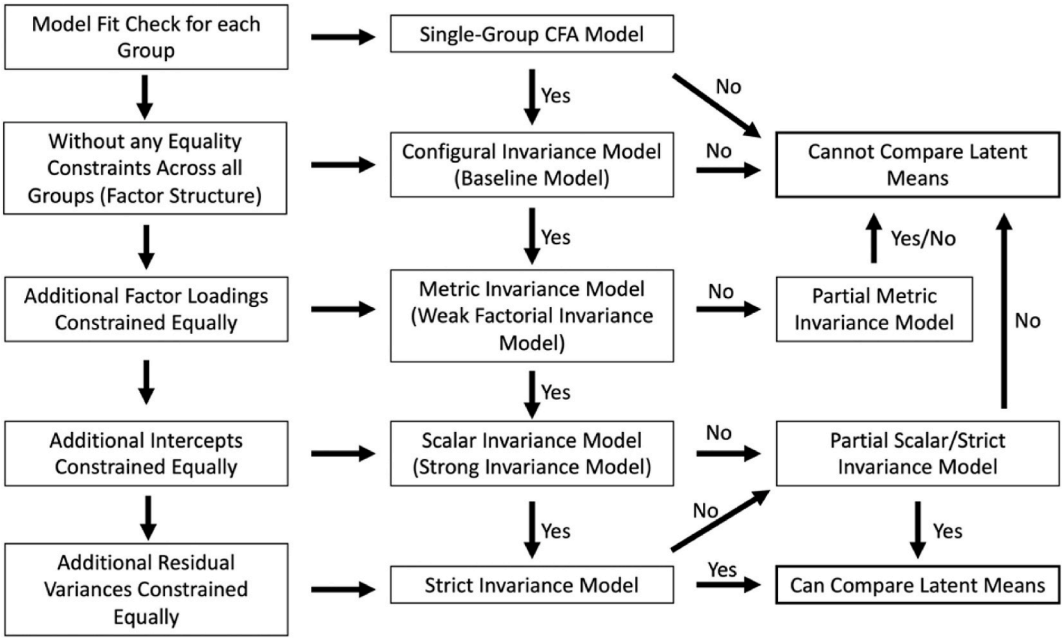

Fig. 1. The procedure and techniques of measurement invariance.

statistics was adopted.

In addition, as the traditional chi-square difference test cannot be used for MLR estimation, the Satorra–Bentler scaled chi-square difference test [60] was introduced. The calculation equation is listed below, in which, C_0 and C_1 represent the scaling correction factor for the null model and alternative model, respectively; D_0 and D_1 are their degrees of freedom (*df*); X_0 and X_1 indicate that their Satorra–Bentler adjusted chi-square values; and T represents the chi-square with the *df* difference of D_0 and D_1. After calculation, the significance of the p-value (CI = 95%) for the distributed T can be checked in the chi-square table.

$$T = \frac{(X_0{*}C_0 - X_1{*}C_1){*}(D_0 - D_1)}{(D_0{*}C_0 - D_1{*}C_1)}$$

The model fit for CFA can be shown by the Root Mean Square Error of the Approximation (RMSEA) [61], the Standardized Root Mean Square Residual (SRMR) [62], the Comparative Fit Index (CFI) [63], and the Tucker–Lewis index (TLI) [64,65] with the acceptance criteria of CFI and TLI >0.90, and RMSEA and SRMR <0.08 [66]. The MI test can begin once this fit has been made. Otherwise, it is not permitted to compare hidden means. However, the chi-square test for the model comparisons is sensitive to the sample size, which could result in bias. Thus, the insensitive ΔCFI, ΔRMSEA, and ΔSRMR will also be calculated in the MI phase of this study. The criteria for these values are ΔCFI ≤0.01, ΔRMSEA ≤0.015, and ΔSRMR ≤0.03 for the metric invariance test, and ΔCFI ≤0.01, ΔRMSEA ≤0.015, and ΔSRMR ≤0.01 for scalar and strict invariance test [67].

The configural model which is the baseline model without any equality constraints to compare the factor structure fit simultaneously across groups rather than single-group CFA is the initial MI model that needs to be evaluated. The metric model can be used if the goodness of fit is satisfactory; otherwise, the comparison of the latent means is not supported.

The metric (weak) invariance model requires constraining factor loadings of each indicator across groups equally, which is afterward to make a comparison with the configural model. The full metric invariance model will be developed if the p-value of the chi-square is non-significant. If not, a partial metric model with loading restrictions can be tried. The weak invariance paradigm, however, is not acceptable for latent means comparison. Thus, whether a partial metric invariance model was built or not, it cannot be used for further latent comparisons [68,69].

The scalar (strong) invariance model, with the equally constrained intercepts based on the full metric model, can be evaluated by comparisons with the full metric invariance model. If the difference is insignificant, the full scalar invariance model can be reached, which leads to two options: checking the strict invariance or performing the comparison of the latent means, for which the scalar invariance model serves as the foundation [68,69]. However, if the difference is significant, the partial scalar model can be constructed by freeing some variant indicators. If the partial scalar was achieved, the comparison of the mean can also be manipulated [70,71].

The (partial) strict invariance model is not often assessed in studies since it requires constraining residual variances equally on the basis of the full scalar invariance model, which is not easy to reach and is not the premise to compare latent means [68,72]. Accordingly, the strict invariance model will not be considered in this study.

The latent mean difference can be explored with at least two scalar invariant indicators in each factor [72]. In this study, the latent means were compared using the partial scalar invariance model, with boys' zero means serving as the reference group.

4.3.2. Moderated mediating effects analysis

On the basis of the invariance model and latent mean comparison results, the further influence of gender difference on paths will be assessed to explore the moderator role of gender in the relationships among the components of CTD. The foundation of a relationship with moderated mediating effects is the existence of a mediating effect first, followed by the moderator's ability to change the routes in this relationship. Thus, the first exploration is whether the mediating effect exists. If so, the moderation role will be followed. Thus, in this study, based on previous studies, we hypothesized that self-efficacy will be the mediator (Me), instant judgment (X) is the independent latent variable, and habitual truth-digging (Y) will be the dependent latent variable in the mediation investigation process. If this process is established, the gender (Mo) will be set as the moderator in these three paths as shown in Fig. 2. This process will be conducted with the Partial Least Squares Structural Equation Modeling (PLS-SEM) via STATA 17.0.

For the mediating effect investigation, Baron and Kenny's tests [73], and Sobel's tests [74] are the traditional ones. Baron and Kenny suggested testing the following three regressions (equations 1–3). They are as follows: the regression of Me on X (1), the direct regression of Y on X without Me as a mediator (2), and the regression of Y on X and Me (3). The mediation effect will be achieved if all three regressions are significant. Accordingly, Sobel's z-score test was introduced to test the significance of the indirect path a × b (equation 4).

$$Me = \mu_1 + aX + e_1 \tag{1}$$

$$Y = \mu_2 + c^{'}X + e_2 \tag{2}$$

$$Y = \mu_3 + cX + bMe + e_3 \tag{3}$$

$$z = \frac{\text{a*b}}{\sqrt{a^2 * S_b^2 + b^2 * S_a^2}} \tag{4}$$

However, Zhao, Lynch Jr [75] concluded that these two methods were not exact and powerful compared to the bootstrap approach by Preacher and Hayes [76]. Thus, to provide a comprehensive picture, this study conducted all three types of investigations on mediating effect, and the mediation will be evaluated using the tree diagram instructions [75] together with statistical results, shown in Fig. 3, including the procedure (a) and (b).

The moderation exploration will be established based on the mediating model. When a third variable has a considerable influence on the relationship between the predictor and the outcome, this third variable is referred to as the moderator [77]. In the moderated mediating effect model like Fig. 2, the moderator (Mo) may affect the regression paths *a*, *b*, or *c*, separately, partly, or simultaneously. The relevance of the interactive relationship between the predictor and the moderator will be examined in the analysis of the continuous moderator using the equations mentioned below (equations 5–7). If the coefficient A, B, or C is significant, then moderation will be established.

$$Me = \mu_1 + aX + a'Mo + AXMo + e_1 \tag{5}$$

$$Y = \mu_2 + bMe + b'Mo + BMeMo + e_2 \tag{6}$$

$$Y = \mu_3 + cX + c^{'}Mo + CXMo + e_3 \tag{7}$$

However, in this study, the moderator is the dichotomous variable. As a result, we directly included the group variable in the PLS-SEM model to examine the significance of the difference in the moderating effects between boys and girls on each path. If the difference in the effects across gender significant, then the moderation can be concluded.

5. Results

5.1. Reliabilities and the baseline model fit

The good reliabilities of the scale in this study are the overall Cronbach alpha = .826, Cronbach $\text{alpha}_{\text{boy}}$ = .861, and Cronbach

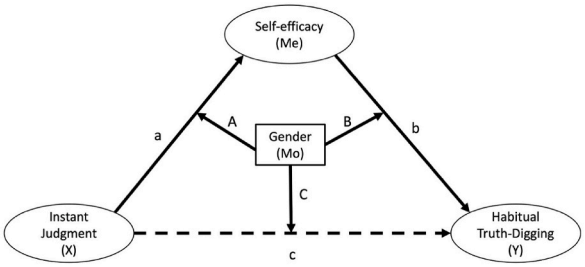

Fig. 2. The hypothesis model of the moderated mediating effects of gender.

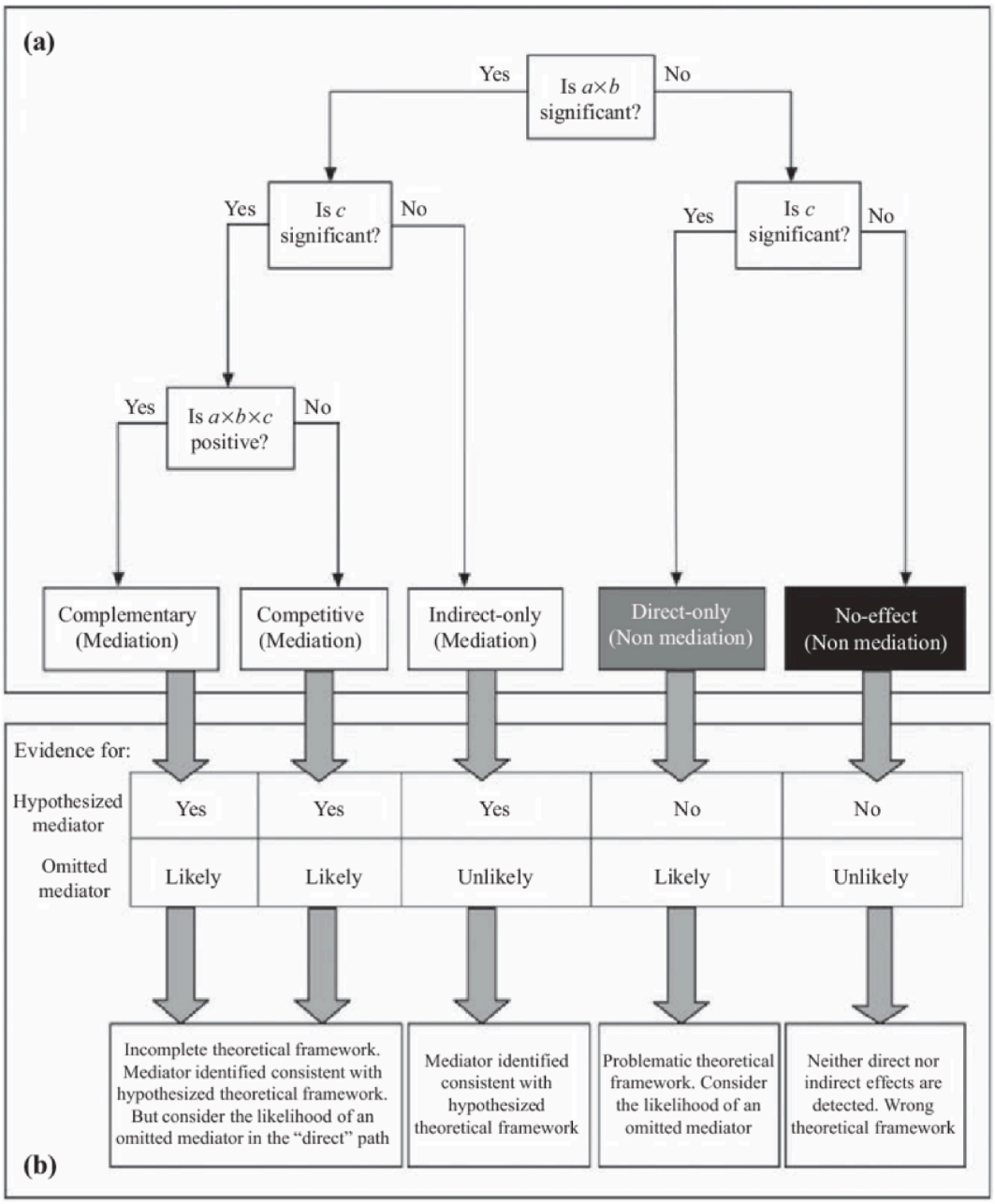

Fig. 3. Building mediation classifications and implications for theory [75].

$alpha_{girl}$ = .777. Table 2 shows the factor loadings, descriptive statistics of items, and factor correlations. All factor loadings were significantly acceptable, and the correlations between instant judgment and self-efficacy and between self-efficacy and habitual truth-digging were significant, but the correlation between instant judgment and habitual truth-digging was not significant, indicating that self-efficacy acted as a mediator between them. The acceptable overall model fit and model fit among gender are listed in Table 3 with CFI and TLI >0.90, and RMSEA and SRMR <0.08, which is the basis for further analysis.

5.2. *Measurement invariance and latent mean comparisons*

Table 4 lists the results of the MI test on comparisons between configural, metric, scalar, and partial scalar models divided by gender.

The model fit of the configural (baseline) model (Model 1) with no constraints was acceptable, which indicated that the configural

Table 2
Factor loadings and descriptive statistics of items in the baseline model.

Items	Standardized Coefficients		Descriptive Statistics			
	Boys (N = 328)	Girls (N = 333)	Boys		Girls	
			Mean	SD	Mean	SD
A1	.601***	.470***	4.00	1.72	4.29	1.66
A2	.706***	.691***	3.69	1.73	3.96	1.78
A3	.659***	.684***	4.04	1.68	4.35	1.77
A4	.597***	.527***	4.23	1.64	4.57	1.61
A5	.689***	.531***	3.70	1.68	4.37	1.68
B1	.830***	.785***	2.77	1.81	2.23	1.66
B2	.756***	.776***	3.25	1.68	2.76	1.82
B3	.657***	.741***	3.31	1.89	2.70	1.88
C1	.758***	.602***	5.30	1.56	5.68	1.20
C2	.724***	.522***	5.01	1.51	4.98	1.47
C3	.719***	.556***	4.95	1.55	5.10	1.41
C4	.747***	.555***	5.12	1.42	5.34	1.22
C5	.598***	.433***	5.04	1.54	5.26	1.41
C6	.632***	.460***	4.78	1.50	4.82	1.53
C7	.736***	.568***	4.97	1.45	4.97	1.44
C8	.665***	.460***	4.85	1.51	4.65	1.48
C9	.696***	.552***	4.90	1.50	5.25	1.28
C10	.587***	.548***	4.92	1.50	5.20	1.34
C11	.672***	.485***	5.12	1.39	5.41	1.30
INSJ with SF	.811***	.614***				
SF with HTD	.175**	.137*				
INSJ with HTD	.017	−.038				

Notes: ***p < .001, **p < .01, *p < .05, confidence interval (CI) = 95%, INSJ = Instant Judgment, SF = Self-Efficacy, HTD = Habitual Truth-Digging.

Table 3
Overall and single-group CFA model fit.

Model	χ2	df	SCF	CFI	TLI	RMSEA	SRMR
Overall	341.988***	149	1.3197	0.935	0.925	0.044	0.048
Boys	290.015***	149	1.4007	0.924	0.913	0.054	0.055
Girls	256.540***	149	1.1666	0.910	0.896	0.047	0.060

Notes: ***p < .001, CI = 95%, CFA = confirmatory factor analysis, SCF = Scaling Correction Factor for Maximum Likelihood Robust (MLR).

Table 4
Comparisons by gender among Configural, metric, scalar, and partial scalar models.

Model	χ^2 ($\Delta\chi^2$)	*df*	CFI	TLI	RMSEA	SRMR	ΔCFI	ΔRMSEA	ΔSRMR
1. Configural	549.608***	298	0.919	0.907	0.051	0.058			
2. Metric	574.806***	314	0.916	0.908	0.050	0.061			
1 VS 2	23.209	16					.003	.001	.003
3. Scalar	624.915***	330	0.905	0.901	0.052	0.064			
1 VS 3	81.015***	32					.014	.001	.006
2 VS 3	55.662***	16					.011	.002	.003
4. Partial Scalar	597.121***	327	0.913	0.909	0.050	0.062			
2 VS 4	21.907	13					.003	.000	.001

Notes: ***p < .001, CI = 95%.

MI was achieved.

The model fit of the metric (weak factorial) invariance model (Model 2) was good. The comparisons between the metric and configural model were nonsignificant, and the differences in values (ΔCFI, ΔRMSEA, and ΔSRMR) were all within the criteria, showing that the full metric invariance model was established.

The model fit of the scalar (strong) invariance model (Model 3) is achieved. The Satorra-Bentler scaled chi-square difference test, in contrast, indicated that the difference between scalar and metric models is substantial. Further, ΔCFI was not within the acceptable range, which indicated that the full scalar invariance model cannot be obtained. As a result, a partial scalar invariance model (Model 4) was evaluated for releasing the intercepts of indicators A5, C7, and C8 based on the modification indices. The model fit of the partial model was good enough, the Satorra–Bentler scaled chi-square difference test between the partial scalar and metric model is nonsignificant, and ΔCFI, ΔRMSEA, and ΔSRMR are all within the range, illustrating that the partial scalar invariance model was built, and further group comparisons could be performed.

Based on the partial scalar model with the freedom of intercepts of three items, the comparisons of the three latent means, with the zero mean of boys as the reference, showed that girls have stronger self-efficacy (standardized $mean_{girl}$ difference = 0.289, SE = 0.099, $p < .01$) and habitual truth-digging (standardized $mean_{girl}$ difference = 0.316, SE = 0.100, $p < .01$) disposition than boys, whereas boys have a higher instant judgment (standardized $mean_{girl}$ difference = −0.400, SE = 0.111, $p < .001$).

5.3. Moderated mediating effects

The results of the test for mediating effects (see Fig. 4) showed that the significant path coefficients of the regressions between the mediator L1, independent latent variable L2, and dependent latent variable L3 were found. Meanwhile, the path direction from L2 to L3 is significant with a negative path coefficient and the indirect effect is significant in Table 5, which indicates that there seems to have a partial mediation (competitive mediating effect) of self-efficacy (L1) between instant judgment (L2) and habitual truth-digging (L3).

According to the mediating effect model, only the path L2 to L1 found a significant moderated effect of gender, and the effect of boys is a bit stronger than that of girls (see Fig. 5 and Table 6). This depicts that, if they have the same level of instant judgment, boys will generate more self-efficacy in their minds than girls. In other words, boys may acquire self-confidence more easily than girls on the

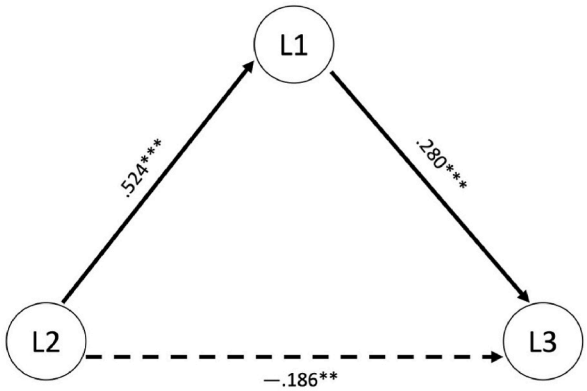

Fig. 4. Mediating effects among the components of CTD.
Note: L1: self-efficacy; L2: instant judgment; L3: habitual truth-digging; $^{**}p < .01$; $^{***}p < .001$.

Table 5
Sobel's testing of indirect effects.

Method	Model Path	Indirect Effect	S.E.	z	95% CI		Direct Effect	Total Effect	RIT	RID
					Low	Up				
Sobel	L2-L1-L3	.147***	.025	5.865	.098	.196	−.186	.039	3.753	.790
Delta		.147***	.025	5.865	.098	.196				
Bootstrap		.147***	.034	4.364	.089	.216				

Note: L1: Self-Efficacy; L2: Instant Judgment; L3: Habitual Truth-Digging; bootstrap replications: 1000; RIT = Ratio of indirect and total effect; RID = Ratio of indirect and direct effect.

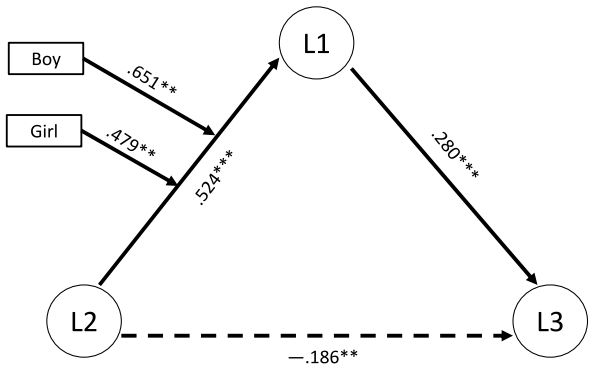

Fig. 5. Moderated mediating effect of gender among the components of CTD.
Note: L1: Self-efficacy; L2: Instant judgment; L3: Habitual truth-digging; $^{**}p < .01$; $^{***}p < .001$.

Table 6
Moderation of gender across relationships of CTD components.

Structural Effect	Global	Boy	Girl	Absolute Difference	*z*	*p*
L2-L1	.524	.651	.479	.172	2.695	.007
L1-L3	.280	.310	.274	.036	.385	.700
L2-L3	−.186	−.181	−.288	.106	1.147	.252

Note: L1: Self-Efficacy; L2: Instant Judgment; L3: Habitual Truth-Digging.

level of instant judgment, indicating that gender moderated the relationships between instant judgment and self-efficacy in the mediating effects model between instant judgment and habitual truth-digging with self-efficacy as the competitive mediator.

6. Discussion and conclusion

The procedure and the aim of this study are to test MI, investigate the latent comparisons across gender, and explore the moderated mediating effects of gender across the components of CTD. The technical explanation of MI on CTD assessment will therefore be covered first. This study is not intended to minimize the value of the studies that do not test MI before latent mean comparisons; rather, it is intended to highlight the significance of MI for determining whether the influence of the scale contributed to the difference to eliminate one influencing factor of gender difference. Thus, for the comparisons among different groups, contexts, and time series, the MI test ahead is a must. Further, from the perspective of the model estimator, some studies adopted the default ML estimator without testing the normality of data, which is not acceptable since the underlined logic of MI is accurate for the estimation of normal distribution data. As a result, the normality test step can be conveniently skipped when MLR is implemented. It should also be noted that the Satorra-Bentler scaled chi-square difference test must be used in place of the conventional chi-square difference test while performing the MLR estimation.

Furthermore, for the partial scalar invariance model, the three items were further analyzed on the reason why they were variant across gender, for example, item A5 "*When making decisions, I tend to be impulsive*." Item C7 "*I like to deal with problems logically (step by step or with a good plan)*." Item C8 "*When judging others' statements, I tend to use reasonable evidence or criteria rather than basing them on my own experience.*" These items tend to be either impulsive or mildly temperamental, which may be noticed as a difference between most girls and boys. It has been discovered that these traits are relatively extreme. Thus, the significant variance can be understandable.

As a result, for the latent mean comparison across gender, it is acceptable to acknowledge that boys and girls have different personalities in some educational contexts because they differ from one another in terms of both physiological and psychological phenomena. Although some studies (e.g., [32–36]) showed no difference in CTD across gender, most studies (e.g., [16,27–29]) and the only one study with MI test [10] reached to the agreement that females scored higher than males in overall CTD and subdimensions. Thus, even though this study was the first to test the MI of 2ES-CTDI, the same tendency was generated as previous studies, which is also, to some extent, a kind of evidence of the validity of this scale. Furthermore, a plausible explanation for the findings in this study that girls have a stronger disposition in self-efficacy and habitual truth-digging while having a lower disposition in instant judgment could be the reason that the characteristics of girls, in general, are dedication, carefulness, and cautiousness. These can direct self-confidence and efficacy, which explains why boys are, somewhat, impulsive, more active, and do not pay more attention to details, such that they have a higher instant judgment disposition. The same results, as Lighthall, Sakaki [78] generated, that boys showed fast decision speed under judgment stress. In addition, Reiter [79] found that when facing a familiar situation, boys will make decisions more quickly than girls, whereas, in an unfamiliar environment, no difference will be shown. Besides, men seem to be overconfident and make decisions fast and alone, while women are sensitive and make decisions in groups [80]. In general, males have stronger decision-making than females [81]. However, the real explanations should be explored in-depth on the brain mechanism across gender in future studies.

The results of the moderated mediating effects also showed a similar tendency of latent difference. First, from the overall perspective, this study got the same trends in the relationships among instant judgment, self-efficacy, and habitual truth-digging. Self-efficacy, a certain kind of inner self-confidence, is often stimulated by self-satisfaction which can be cumulated by instant judgment positively and vice versa [14,82]. Self-efficacy is positively connected with routine truth-seeking as a motivator, and higher self-efficacy may foster a high-level tendency to seek the truth [83]. Habitual truth-digging is the process of seeking truth, including asking questions, finding evidence, analyzing step by step, and evaluating, which is a "time-consuming" procedure, and, theoretically, will be contrary to instant judgment, and fast decision-making. Kahneman [84] concluded that there are both fast and slow thinking in human-mind, and they fight with each other to get the balance, which is also shown in this study that there was a negative relationship between instant judgment and habitual truth-digging with self-efficacy as the competitive mediator.

Backing to the moderator of gender, based on the results of latent mean differences that discussed the cut-off points of gender difference on each factor, the moderating effects discussed the influence on the relationship paths; however, there may be some potential correlations between the latent difference and moderation process. This study discovered no gender moderation in the links between instant judgment and habitual truth-digging, and between self-efficacy and habitual truth-digging, but in the path of instant judgment and self-efficacy, and the gender differences in each factor. No such previous studies can be found for these non-significant moderation effects, but the gender differences discussions have been stated in previous studies like what this study discussed before. When the gender impacts on the links between instant judgment and habitual truth-digging and self-efficacy and truth-digging are

combined for discussion, boys may acquire more self-efficacy through the degree of instant judgment. However, girls have a stronger disposition for self-efficacy and habitual truth-digging while lower in instant judgment. The same significant gender influence can be found as Lighthall, Sakaki [78] showed that boys easily decided fast leading to sharp increasing satisfaction, which has given evidence of the stronger moderating effect of boys on this path. However, despite having a higher level of instant judgment and a steeper growing slope, boys have lower self-efficacy than girls, which can be explained by the fact that boys can readily get and dismiss self-efficacy, proving the prior result that they are impulsive.

Limitations still exist in this study. Although the partial scalar invariance model was achieved, the comparison of the latent means and the moderated mediating effects showed plausible results, and the full scalar invariance model and the full strict model could not be established. Additionally, the comparisons of subdimensions of CTD across gender were conducted based on 2ES-CTDI. Varying scales may attain different levels of MI, contain different latent components, and produce different findings. Moreover, although the theoretical framework of this scale dividing CTD into instant judgment, self-efficacy, and habitual truth-digging was proved stable, this framework was newly built, which needs more future studies to explore this theory in different contexts. Furthermore, while the results showed that the cause of the distinct tendencies of boys and girls in CTD is not the scale but the gender difference in personality, the samples are all Chinese from southwestern China, and the disposition may vary between cultures. Finally, this study aims to figure out the MI, latent means comparison, and gender moderation, and the in-depth explanations of the reason why the differences existed across gender cannot be given.

Similar to earlier research, this study found that girls fared better than boys on CTD and several subdimensions. However, some contradictory results have been found in prior studies, which could be attributed to the diverse student backgrounds in terms of nations, majors, and instruments. Thus, in addition to considering the natural distinction between boys and girls, the future educational practice of assessing CTD for undergraduates should consider both these results and their situations to analyze if there is MI in their scale and if there is a gender difference in their context, and then take the correspondent actions that are suitable for their students. Furthermore, the three variant items can be evaluated in future studies to see if they are invariant, or they may be eliminated in subsequent studies to avoid the variance when comparing the gender difference in latent mean level.

The directions for instructional focus on the assessment and cultivation of undergraduate students' CTD are: (a) gender differences and the moderating effect should be paid more attention to. Since boys' CTD self-efficacy and habitual truth-digging levels are lower, instructors can maintain their self-efficacy by encouraging or praising and cultivating their patience to solve complex problems logically and persistently. (b) While together with mediating effect results, self-efficacy has competitive effects on the relationship between instant judgment and habitual truth-digging, indicating more instant judgment and self-efficacy may decrease habitual truth-digging. Thus, teachers should monitor this frequently to find a balance among them for the promotion of students' CTD.

To conclude, the 2ES-CTDI has a partial scalar function on the equality of CTD assessment across gender, which provides evidence that the gender difference among the latent factors (self-efficacy, instant judgment, and habitual truth-digging) and the components of CTD is not caused by the scale but by the different personality inclination of boys and girls. In addition, self-efficacy partially mediated the relationships between instant judgment and habitual truth-digging, and boys may acquire self-confidence more easily than girls on the level of instant judgment, necessitating educators to pay more attention to this distinction in gender in their educational practice on undergraduates' CTD propensity.

Appendix. Employer-Employee-Supported Critical Thinking Disposition Inventory (2ES-CTDI).

Author contribution statement

Yong Liu: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Attila Pásztor: Supervised the experiments; Funding acquisition; Wrote the paper for reviewing and editing the revision.

Funding statement

This work is supported by China Scholarship Council (Grant Number: 202008500108), Tempus Public Foundation of Hungary Government (Grant Number: SHE-16610-004/2020), the Research Program for Public Education Development of the Hungarian Academy of Sciences (Project Number: KOZOKT2021-16, and University of Szeged (Open Access Grant Number: 5885).

Data availability statement

Data will be made available on request.

Declaration of interest's statement

The authors declare no competing interests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e14664.

References

[1] R.H. Ennis, Critical thinking: reflection and perspective, in: Inquiry: Critical Thinking across the Disciplines, Prentice-Hall, Upper Saddle River, New Jersey, 2011, pp. 4–18.
[2] S.P. Norris, The meaning of critical thinking test performance: the effects of abilities and dispositions on scores, in: Critical Thinking and Reasoning: Current Research, Theory, and Practice, Hampton Press, Inc, Cresskinll, 2003, pp. 315–329. F.D. J.
[3] D.N. Perkins, E. Jay, S. Tishman, Beyond abilities-A dispositional theory of thinking, Merrill-Palmer Quarterly-Journal of Developmental Psychology 39 (1) (1993) 1–21.
[4] P. Facione, Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction—The Delphi Report, California Academic Press, Millbrae, CA, 1990.
[5] P.A. Facione, N.C. Facione, California Critical Thinking Disposition Inventory, California Academic Press, Millbrae, CA, 1992.
[6] P.R. Pintrich, et al., Reliability and predictive validity of the motivated strategies for learning questionnaire (MSLQ), Educ. Psychol. Meas. 53 (3) (1993) 801–813.
[7] J.C. Ricketts, The Efficacy of Leadership Development, Critical Thinking Dispositions, and Student Academic Performance on the Critical Thinking Skills of Selected Youth Leaders, University of Florida, 2003.
[8] J. Yoon, Development of an Instrument for the Measurement of Critical Thinking Disposition in Nursing, The Catholic University of Education, Seoul, Korea, 2004, p. 69.
[9] E.M. Sosu, The development and psychometric validation of a critical thinking disposition scale, Think. Skills Creativ. 9 (2013) 107–119.
[10] M.J. Bravo, et al., An adaptation of the critical thinking disposition scale in Spanish youth, Think. Skills Creativ. 38 (2020), 100748.
[11] S. Quinn, et al., Development and validation of the student-educator negotiated critical thinking dispositions scale (SENCTDS), Think. Skills Creativ. 38 (2020), 100710.
[12] Y. Liu, A. Pasztor, Design and validate the employer-employee-supported critical thinking disposition inventory (2ES-CTDI) for undergraduates, Think. Skills Creativ. 46 (2022), 101169.
[13] A. Bandura, F.J. Jourden, Self-regulatory mechanisms governing the impact of social comparison on complex decision making, J. Pers. Soc. Psychol. 60 (6) (1991) 941.
[14] B.B. Leslie, J. Moilanen, Advancing self-efficacy and intuitive decision-making calls to action, in: Adult Education Research Conference, 2010.
[15] P.A. Facione, N.C. Facione, C.A. Giancarlo, Setting expectations for student learning: new directions for higher education, in: Millbrae: Millbrae, California Academic Press, 1997.
[16] C.M. Walsh, R.C. Hardy, Dispositional differences in critical thinking related to gender and academic major, J. Nurs. Educ. 38 (4) (1999) 149–155.
[17] N. Yenice, Investigating pre-service science teachers' critical thinking dispositions and problem solving skills in terms of different variables, Educ. Res. Rev. 6 (6) (2011) 497–508. https://www.scopus.com/inward/record.uri?eid=2-s2.0-79961099501&partnerID=40&md5=9844893ea32bc3e1147e6ec13e3ffe06.
[18] A. Elkoca, Critical thinking skills of nursing candidates, Int. J. Caring Sci. 12 (3) (2019) 1600–1606. https://internationaljournalofcaringsciences.org/docs/30_ozcan_original_12_3.pdf.
[19] A.C. Biber, A. Tuna, L. Incikabi, An investigation of critical thinking dispositions of mathematics teacher candidates, Educ. Res. 4 (2013) 109–117. https://www.interesjournals.org/articles/an-investigation-of-critical-thinking-dispositions-of-mathematics-teacher-candidates.pdf.
[20] K. Bilen, O. Ercan, E. Akcaozoglu, Identification of critical thinking dispositions of teacher candidates, Anthropol. 16 (3) (2013) 435–448.
[21] X. Du, et al., PBL and critical thinking disposition in Chinese medical students–A randomized cross-sectional study, Journal of Problem Based Learning in Higher Education 1 (1) (2013) 72–83.
[22] A.I. Mwalongo, Gender differences in the use of asynchronous discussion forums and quizzes for promoting critical thinking dispositions, in: Gender and Diversity: Concepts, Methodologies, Tools, and Applications, 2018, pp. 476–492.
[23] E. Demirhan, A.N. Koklukaya, The critical thinking dispositions of prospective science teachers, Procedia-Social and Behavioral Sciences 116 (2014) 1551–1555.
[24] H. Fitriani, et al., Critical thinking disposition of prospective science teachers at IKIP mataram, Indonesia, J. Phys. Conf. 1108 (2018), 012091.
[25] E.M. Whitney, J. Aleksejuniene, J.N. Walton, Critical thinking disposition and skills in dental students: development and relationship to academic outcomes, J. Dent. Educ. 80 (8) (2016) 948–958. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84980319541&partnerID=40&md5=e487a2f54c6f712cb8c8d88d66e6f579.
[26] X. Ding, The effect of WeChat-assisted problem-based learning on the critical thinking disposition of EFL learners, Int. J. Emerg. Technol. Learn 11 (12) (2016).
[27] T.H. Bers, M. McGowan, A. Rubin, The disposition to think critically among community college students: the California critical thinking dispositions inventory, J. Gen. Educ. 45 (3) (1996) 197–223.
[28] M. Kirbaslar, Z. Ozsoy-Gunes, The effect of critical thinking disposition on entrepreneurship levels: a study on future teachers, Procedia Soc. Behav. Sci. 174 (2015) 199–207.
[29] N. Li, Reasonable or unwarranted? Benevolent gender prejudice in education in China, Asia-Pac. Educ. Res. 31 (2) (2022) 155–163.
[30] M. Najafi, et al., Trend of tendency to critical thinking among medical students in Tehran University of Medical Sciences, 2010-2015: a longitudinal study, J. Educ. Health Promot. 11 (1) (2022) 29.
[31] S.K. Dennett, M.A. DeDonno, A comparison between Chinese and American male and female college students' critical thinking dispositions, Int. J. Chin. Educ. 10 (3) (2021).
[32] Z. Koyunlu Unlu, I. Dokme, Science teacher candidates' epistemological beliefs and critical thinking disposition, Egit. Arastirmalari - Eurasian J. Educ. Res. 2017 (72) (2017) 203–220.
[33] B. Demirbag, A. Unisen, A. Yesilyurt, Training of critical thinking skills in teacher candidates and placebo effect: a quasi-experimental study, Egit. Arastirmalari - Eurasian J. Educ. Res. (63) (2016) 375–392.
[34] O. Ozyurt, Examining the critical thinking dispositions and the problem solving skills of computer engineering students, Eurasia J. Math. Sci. Technol. Educ. 11 (2) (2015) 353–361.
[35] I. Kilic, T. Yazici, S.I. Topalak, Critical thinking disposition of music teachers, Egit. Arastirmalari - Eurasian J. Educ. Res. 2017 (72) (2017) 185–202.
[36] Y. Liu, A. Pasztor, Effects of problem-based learning instructional intervention on critical thinking in higher education: a meta-analysis, Think. Skills Creativ. 45 (2022), 101069.
[37] N. Shida, et al., Critical thinking dispositions among polytechnic students: why does it matter? Int. J. Eng. Technol. 7 (3) (2018) 357–361.
[38] N. Mousazadeh, et al., Effective factors in critical thinking disposition in nursing students, Educ. Res. Int. 2021 (2021), 5580010.
[39] G. Park, Influence of critical thinking disposition, nursing professional values and satisfaction in clinical practice of nursing students, Asia Life Sci. 20 (2) (2019) 37–49. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85084683098&partnerID=40&md5=187256279fc276534d62eb9011a2ca0e.
[40] R.H. Stupnisky, et al., The interrelation of first-year college students' critical thinking disposition, perceived academic control, and academic achievement, Res. High. Educ. 49 (6) (2008) 513.
[41] P. Álvarez-Huerta, A. Muela, I. Larrea, Disposition towards Critical Thinking and Student Engagement in Higher Education, Innovative Higher Education, 2022.
[42] Y. Liu, A. Pasztor, Impacts of demographic factors on critical thinking disposition among undergraduates, in: J.B. Fejes, A. Pásztor-Kovács (Eds.), 18th Conference on Educational Assessment, University of Szeged, Szeged, Hungary, 2022.
[43] B. Dokmecioglu, Y. Tas, S. Yerdelen, Predicting students' critical thinking dispositions in science through their perceptions of constructivist learning environments and metacognitive self-regulation strategies: a mediation analysis, Educ. Stud. (2020) 1–18.
[44] J.-a. Kim, E.-H. Seo, C.-J. Kim, Mediating effect of major satisfaction on the influence of critical thinking disposition on disaster recognition, Int. J. Adv. Cult. Technol. 6 (1) (2018) 23–31.

[45] M.H.M. Cheng, Z.H. Wan, Exploring the effects of classroom learning environment on critical thinking skills and disposition: a study of Hong Kong 12th graders in Liberal Studies, Think. Skills Creativ. 24 (2017) 152–163.
[46] Y. Song, Y. Lee, J. Lee, Mediating effects of self-directed learning on the relationship between critical thinking and problem-solving in student nurses attending online classes: a cross-sectional descriptive study, Nurse Educ. Today 109 (2022), 105227.
[47] A. Orhan, California critical thinking disposition inventory: reliability generalization meta-analysis, J. Psychoeduc. Assess. 40 (2) (2022) 202–220.
[48] S. Sk, S. Halder, Critical thinking disposition of undergraduate students in relation to emotional intelligence: gender as a moderator, Heliyon 6 (11) (2020), e05477.
[49] G. Ali, R. Awan, Thinking based instructional practices and academic achievement of undergraduate science students: exploring the role of critical thinking skills and dispositions, J. Innovative Sci. 7 (1) (2021) 56–70.
[50] C.J. Fong, et al., A meta-analysis on critical thinking and community college student achievement, Think. Skills Creativ. 26 (2017) 71–83.
[51] G.C. Abiogu, et al., Cognitive-behavioural reflective training for improving critical thinking disposition of nursing students, Medicine 99 (46) (2020) 12.
[52] A.R. Damasio, A second chance for emotion, Cognitive Neuroscience of Emotion 12 (2000) 23.
[53] N.L. Stedman, A.C. Andenoro, Identification of relationships between emotional intelligence skill and critical thinking disposition in undergraduate leadership students, Journal of Leadership Education 6 (1) (2007) 190–208.
[54] K.G. Joreskog, A.S. Goldberger, Estimation of a model with multiple indicators and multiple causes of a single latent variable, J. Am. Stat. Assoc. 70 (351) (1975) 631–639.
[55] D. Sorbom, A general method for studying differences in factor means and factor structure between groups, Br. J. Math. Stat. Psychol. 27 (2) (1974) 229–239.
[56] Y. Xu, Assessing Measurement Invariance and Latent Mean Differences with Bifactor Multidimensional Data in Structural Equation Modeling, Arizona State University, 2018, p. 99.
[57] F. Chen, What happens if we compare chopsticks with forks? The impact of making inappropriate comparisons in cross-cultural research, J. Pers. Soc. Psychol. 95 (2008) 1005–1018.
[58] C.-H. Li, The Performance of MLR, USLMV, and WLSMV Estimation in Structural Regression Models with Ordinal Variables, Michigan State University, 2014, p. 148.
[59] L.K. Muthen, B.O. Muthen, Mplus User's Guide, 7 ed., Muthén & Muthén, Los Angeles, CA: Los Angeles, CA, 1998-2015.
[60] A. Satorra, P.M. Bentler, Ensuring positiveness of the scaled difference chi-square test statistic, Psychometrika 75 (2) (2010) 243–248.
[61] J.H. Steiger, Structural model evaluation and modification: an interval estimation approach, Multivariate Behav. Res. 25 (2) (1990) 173–180.
[62] P.M. Bentler, EQS Structural Equations Program Manual, vol. 6, Multivariate software Encino, CA, 1995.
[63] P.M. Bentler, Comparative fit indexes in structural models, Psychol. Bull. 107 (2) (1990) 238–246.
[64] P.M. Bentler, D.G. Bonett, Significance tests and goodness of fit in the analysis of covariance structures, Psychol. Bull. 88 (3) (1980) 588.
[65] L.R. Tucker, C. Lewis, A reliability coefficient for maximum likelihood factor analysis, Psychometrika 38 (1) (1973) 1–10.
[66] L.t. Hu, P.M. Bentler, Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives, Struct. Equ. Model.: A Multidiscip. J. 6 (1) (1999) 1–55.
[67] F.F. Chen, Sensitivity of goodness of fit indexes to lack of measurement invariance, Struct. Equ. Model.: A Multidiscip. J. 14 (3) (2007) 464–504.
[68] K.A. Bollen, Structural Equations with Latent Variables, vol. 210, John Wiley & Sons, 1989.
[69] J.L. Horn, J.J. McArdle, A practical and theoretical guide to measurement invariance in aging research, Exp. Aging Res. 18 (3) (1992) 117–144.
[70] H.W. Marsh, D. Hocevar, Application of confirmatory factor analysis to the study of self-concept: first- and higher order factor models and their invariance across groups, Psychol. Bull. 97 (1985) 562–582.
[71] A.C. Carle, R.E. Millsap, D.A. Cole, Measurement bias across gender on the Children's Depression Inventory: evidence for invariance from two latent variable models, Educ. Psychol. Meas. 68 (2) (2008) 281–303.
[72] B.M. Byrne, R.J. Shavelson, B. Muthen, Testing for the equivalence of factor covariance and mean structures: the issue of partial measurement invariance, Psychol. Bull. 105 (3) (1989) 456–466.
[73] R.M. Baron, D.A. Kenny, The moderator–mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations, J. Pers. Soc. Psychol. 51 (6) (1986) 1173.
[74] M.E. Sobel, Asymptotic confidence intervals for indirect effects in structural equation models, Socio. Methodol. 13 (1982) 290–312.
[75] X. Zhao, J.G. Lynch Jr., Q. Chen, Reconsidering Baron and Kenny: myths and truths about mediation analysis, J. Consum. Res. 37 (2) (2010) 197–206.
[76] K.J. Preacher, A.F. Hayes, SPSS and SAS procedures for estimating indirect effects in simple mediation models, Behav. Res. Methods Instrum. Comput. 36 (4) (2004) 717–731.
[77] J. Cohen, et al., Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences, 3 ed., Routledge, 2002.
[78] N.R. Lighthall, et al., Gender differences in reward-related decision processing under stress, Soc. Cognit. Affect Neurosci. 7 (4) (2012) 476–484.
[79] K.R. Reiter, Gender Differences in Decision Making when Faced with Multiple Options, 2013. https://digitalcommons.csbsju.edu/cgi/viewcontent.cgi?article=1010&context=psychology_students.
[80] G. Tovmasyan, Gender differences in decision-making and leadership: evidence from Armenia, Business Ethics and Leadership 4 (2020).
[81] V. Singh, et al., Examining cognitive sex differences in elite math intensive education: preliminary evidence from a gender inequitable country, Trends in Neuroscience and Education 26 (2022), 100172.
[82] R. Lugo, et al., The moderating influence of self-efficacy on interoceptive ability and counterintuitive decision making in officer cadets, Journal of Military Studies 7 (2017) 1–9.
[83] L. Huang, et al., General self-efficacy mediates the effect of family socioeconomic status on critical thinking in Chinese medical students, Front. Psychol. 9 (2019).
[84] D. Kahneman, Thinking, Fast and Slow, Macmillan, 2011.