MALAYSIAN ONLINE JOURNAL OF PSYCHOLOGY & COUNSELING

CRITICAL THINKING DISPOSITION AMONG PAKISTANI UNIVERSITY STUDENTS: PSYCHOMETRIC PROFILES FOR YOON'S CRITICAL THINKING DISPOSITION INSTRUMENT

Lubna Shaheen[1], Harris Shah Abd Hamid[2] & Siaw Yan Li[1]*

ABSTRACT

This research aims to validate Yoon's Critical Thinking Disposition (YCTD) instrument in the context of Pakistan by examining its cross-sectional and construct validity, as well as its measurement invariance across different subgroups (male/female). The study involved a sample of 390 postgraduate university students from Pakistan. The findings were compared to previous research efforts conducted by Shin et al. (2015), Yang et al. (2009), and Yoon (2008). The results confirmed the validity of the seven-factor model of the YCTD and provided evidence of its reliability, as indicated by Cronbach's alpha values exceeding 0.50. The item composite reliability values further supported the validity of the instrument. The factor loadings and significant differences in correlation coefficients between the subscales provided strong evidence for the distinctiveness of the seven subscales. The measurement invariance analyses demonstrated that the instrument is invariant across different genders in the context of Pakistan, indicating its suitability for comparing critical thinking disposition levels. Additionally, the study emphasized the importance of promoting and measuring critical thinking in the education system, with the YCTD serving as a valuable tool for evaluating interventions and monitoring changes in critical thinking disposition over time. The findings have implications for educators and policymakers, highlighting the need for cultivating critical thinking skills in diverse groups and guiding individual-level interventions based on the specific strengths and weaknesses identified by the YCTD. Overall, this study provides strong evidence supporting the validity and reliability of the YCTD instrument in the Pakistani context.

Keywords: *Critical Thinking Disposition, Reliability, Validity, Factor Analysis, Measurement Invariance*

Volume 10 (1), June 2023

[1]Department of Educational Psychology and Counselling, Faculty of Education, Universiti Malaya, Kuala Lumpur, MALAYSIA

[2]Faculty of Management, Education & Humanities, University College MAIWP International, Kuala Lumpur, MALAYSIA

Corresponding Author:
yanli@um.edu.my

INTRODUCTION

The development of critical thinking (CT) has long been considered a priority and agreed-upon key competence of higher education (Bali, 2015; Li *et al.*, 2021). People who have excellent CT skills exhibit keen thoughts, a burning curiosity, and a thirst for reliable information (Papp *et al.*, 2014). Being able to think critically has been shown to be a strong predictor of academic achievement and making successful life decisions in general (Butler *et al.*, 2017; Ren *et al.*, 2020). It is also linked with positive outcomes in education and other facets of intellectual life. There is a significant discussion within the field of CT beginning with disagreements about its specific content and definition, proceeding through the philosophical basis of how CT is approached, to the importance of dispositions to think critically.

Facione (1990) identifies CT as a tool for inquiry. The Delphi Report is a 1990 study by the American Philosophical Association entitled "Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction" defined CT as a purposeful, self-regulatory judgment that results in interpretation, analysis, evaluation, inference, and explanation, as these skills are essential when thinking about how to approach issues, questions, and problems. CT has two involving aspects; a skill dimension and a disposition dimension (Facionc & Facione, 1992; Facione, 1990; Facione, 2000; Facione, 2011; Paul & Binker, 1990; Siegel, 1997). According to Facione (2000), skills and dispositions are two distinct things; they support each other in the process and practice of critical thinking (Figure 1). A strong overall propensity toward critical thinking, for example, is crucial to assuring the usage of CT skills and would motivate an individual to master them (Facione, 2000). Moreover, there are a few different taxonomies for important thinking dispositions, but features like open-mindedness, intellectual curiosity, and reflective thinking are present in all of them (Bravo *et al.*, 2020).

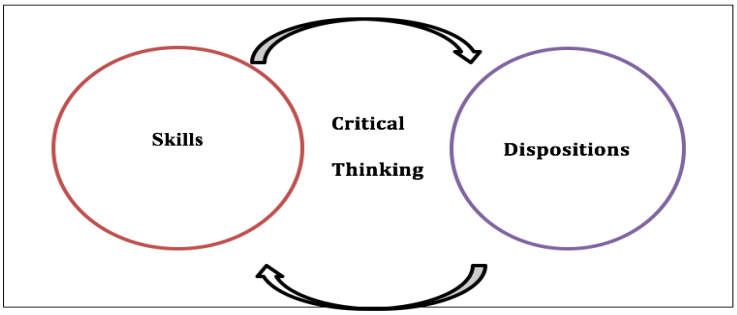

Figure 1. Relationship Between the Skills and the Dispositions of Critical Thinking

Moreover, there have been repeated calls for greater CT ability among university students and current changes in the higher education require strong CT earlier in a students' career (Wolcott & Sargent, 2021). Despite these repeated calls, Pincus *et al.* (2017) pointed out that little educational progress has been made. To encourage and assist in a greater focus on critical thinking, professional organizations are providing competency frameworks and CT resources. Nevertheless, a continuing question is whether educational efforts are sufficient to meet the needs of students and the profession. Besides, questions have arisen about whether CT can be taught and about the effectiveness of educational efforts (Rebele & Pierre, 2019).

There is general agreement in the extant literature that measures of critical thinking should capture and reflect cognitive as well as dispositional components. However, analyses of existing instruments

show a greater emphasis on measurement of cognitive dimensions with little or no consideration given to the dispositional components of critical thinking (Norris, 2003; Poondej & Lerdpornkulrat, 2015; Thomas & Hayes, 2021; Yockey, 2016). Furthermore, several instruments are used to measure critical thinking disposition (CTD) in education. Some of the most significant instruments used for CTD in education include: California Critical Thinking Disposition Inventory (CCTDI) (Facionc & Facione, 1992). This is a well-established instrument that assesses seven critical thinking dispositions, including truth-seeking, open-mindedness, analyticity, systematicity, confidence in reasoning, inquisitiveness, and maturity. Although widely used, this scale is not without its problems. Results from cross-validation studies (İskifoğlu & Ağazade, 2013; Walsh & Hardy, 1997; Walsh *et al.*, 2007; Yeh, 2002) show inconsistencies in the pattern of item loadings, excessive cross loading of items, overlap of constructs, and instability of the hypothesised factor structure, calling into question the validity and reliability for the CCTDI subscales.

Critical Thinking Disposition Scale (Sosu, 2013) is another widely used instrument that measures six dimensions of critical thinking disposition, including inquisitiveness, systematicity, analyticity, truth-seeking, open-mindedness, and confidence in reasoning. Ennis-Weir Critical Thinking Essay Test (Werner, 1991), an essay-based instrument that assesses critical thinking dispositions by measuring a student's ability to analyze, evaluate, and synthesize information. On the other hand, Halpern Critical Thinking Assessment (HCTA) (Halpern, 2010) measures five critical thinking dispositions, including truth-seeking, open-mindedness, analyticity, systematicity, and critical thinking self-confidence. In addition, Cornell Critical Thinking Test (CCTT) (Ennis *et al.*, 2005) is a standardized test that measures critical thinking dispositions by assessing a student's ability to analyze and evaluate information.These instruments are frequently used to assess students' CTD in various educational settings, including K-12 schools, colleges, and universities. By measuring CTD, educators can identify areas for improvement and develop targeted interventions to enhance students' critical thinking skills. However, tools developed from countries other than Pakistan are limited in sensitively capturing the cognitive tendency associated with contextual variations of Pakistani students' perception of every item (Yoon, 2008). With the increasing attention to CT at higher education level, the lack of reliable and valid evaluation tools has been identified as one of the barriers to instruct and evaluate students at university level (Shin *et al.*, 2015). Therefore, these issues have led authors to recommend the validation of Yoon's Critical Thinking Disposition (YCTD) instrument.

YCTD is a theory driven self-rated instrument developed by Yoon (2008) to assess students' critical thinking disposition (CTD). The YCTD focuses on seven areas of CTD: 1) objectivity, 2) prudence, 3) systematicity, 4) intellectual eagerness, 5) intellectual fairness, 6) healthy skepticism, 7) CT self-confidence. Taken together, these seven dimensions consist of disposition for critical thinking (Yoon, 2008). The seven key dimensions of YCTD are based on the California Critical Thinking Disposition Inventory (CCTDI) which was developed by (Facionc & Facione, 1992). In the proposed seven dimensions of YCTD, the objectivity in CT refers to a propensity of rejecting personal biases while systematicity is the tendency of trying to approach challenges/problems in a systematic way. Prudence is defined as the habit of seeing the complexity of issues and intellectual eagerness/curiosity as the habit to learn and know about things. Healthy skepticism refers to the tendency of always seeking the best possible understanding of any given situation while intellectual fairness refers to the propensity of thinking with others' point of view. Finally, the CT self-confidence is identified as the inclination to believe reflective thinking, illuminate problems and form decisions (Shin *et al.*, 2015; Yoon, 2008).

The YCTD instrument has been translated and validated in several studies (Du *et al.*, 2022; Kim, 2012; Kim *et al.*, 2020; Kim, 2010; Qiuhuan, 2008; Shin *et al.*, 2015; Shin *et al.*, 2006; Yang *et al.*, 2009; Yang, 2010). The Cronbach's α of the Chinese version of the YCTD was 0.78 (Qiuhuan, 2008), and Shin *et al.* (2015) reported YCTD reliability using Cronbach α coefficient of 0.84. Kim (2012), reported Cronbach's α reliability of YCTD as 0.86. However, to the best of our knowledge, no study has validated the YCTD instrument in the context of Pakistan. Moreover, Nair and Stamler (2013) have reported an urgent need for examination of the instruments' construct validity. In addition, a valid comparison of self-report instruments such as the YCTD requires that the constructs have a similar meaning across groups (Shin *et al.*, 2015).

OBJECTIVE OF STUDY

The purpose of this study was twofold. Firstly, it aimed to validate the seven-factor model of the YCTD (Yoon's Critical Thinking Disposition) instrument in the context of Pakistani postgraduate (PG) students. The study sought to provide evidence supporting the structural validity of the YCTD by examining whether the seven-factor model accurately represents critical thinking disposition among Pakistani PG students. Secondly, the study aimed to investigate the measurement invariance of the YCTD across gender. By analyzing cross-sectional data, the study sought to compare the patterns of responses to the YCTD and determine if the instrument's measurement properties hold consistently across diverse genders of Pakistani PG students.

METHODOLOGY

Study Design

The study was a validation study of the YCTD using cross-sectional survey design based on a self-reported questionnaire. A cross-sectional survey design was employed because data from a sample of population elements was collected at one point in time (Babbie, 2020).

Participants

The researcher opted to perform the study with postgraduate (PG) students in three public universities of Quetta district, Balochistan, Pakistan. The criteria for inclusion were identified. The inclusion criteria consistently, reliably, uniformly, and objectively identify the study population (Garg, 2016). For the current study, the inclusion criteria were defined as only those students who were able to participate in this study which are:

i. current enrolled PG student from any of three sampled public universities in Quetta district,

ii. PG students only from the department of Education, Sociology, English language and linguistics and Psychology (the rationale to select these departments is that these are the common established departments which can be found in all public universities of Quetta district, Balochistan) (BUITMS, 2019; SBKWU, 2019; UOB, 2019),

iii. already passed their 1st semester at least,

iv. age should not be less than 20 and more than 45,

v. both male/female can participate.

vi.

Based on this inclusion criteria, 390 male and female PG students participated in the present study. The number of students from each university was based on the total number of PG students in the three sampled universities. Figure 2 illustrates the mathematical expression to determine the sample needed from each targeted university.

$$\text{Number of students required from each university} = \frac{\text{Number of students in a university}}{\text{Total number of students}} \times \text{Number of samples}$$

Figure 2. Mathematical Expression to Determine the Sample Needed

Data Collection

At each university, student participants were introduced to the study through a covering letter and were asked to use the YCTD to evaluate their CTD. The covering letter functioned to assure participants of the confidentiality of the data provided. The YCTD questionnaire was administered to PG students during a general lecture of their course. All responses were collected immediately after the lecture. The demographic Table 1 provides an overview of the participants' characteristics in terms of gender, age, faculty, semester, and university. The sample consisted of 205 male participants (52.6%) and 185 female participants (47.4%). This indicates a relatively balanced distribution of gender within the study. The majority of participants fell within the age range of 20-25 years, with 304 participants (77.9%) belonging to this category. A smaller proportion of participants were between the ages of 26-35 years (69 participants, 17.7%), and an even smaller group fell within the age range of 36-40 years (17 participants, 4.4%). The participants represented a diverse range of academic disciplines. The largest group was from the Education faculty with 128 participants (32.8%), followed by Sociology with 94 participants (24.1%). The participants were distributed across different stages of their academic semesters. The highest number of participants were in the 3rd semester with 129 participants (33.1%), followed by the 5th semester and above with 149 participants (38.2%). Moreover, the participants were affiliated with three different universities, denoted as U1, U2, and U3. U2 had the largest representation with 195 participants (50.0%), followed by U3 with 133 participants (34.1%). U1 had the smallest number of participants with 62 (15.9%) (see Table 1). Overall, the sample consisted of a diverse group of participants in terms of gender, age, faculty, semester, and university. This diversity enhances the generalizability of the study's findings and allows for a more comprehensive understanding of critical thinking disposition among Pakistani PG students.

Table 1. Profiles of Respondents according to Demographic Characteristics (N=390)

Variables	Frequency	Percentage
Gender		
Male	205	52.6
Female	185	47.4
Age (years)		
20-25	304	77.9
26-35	69	17.7
36-40	17	4.4
Faculty		
Education	128	32.8
Sociology	94	24.1
English	80	20.5
Psychology	88	22.6
Semester		
2nd	77	19.7
3rd	129	33.1
4th	35	9.0
5th and above	149	38.2
University		
U1	62	15.9
U2	195	50.0
U3	133	34.1

Instrument

The YCTD (Yoon, 2008) was used to measure the participants' levels of CTD in the present study. YCTD consists of 27 items and uses a 5-point Likert rating scale ranging from 1 (strong disagreement) to 5 (strong agreement). The instrument's seven subscales include objectivity, prudence, systematicity, intellectual eagerness/curiosity, intellectual fairness, healthy skepticism, and CT self-confidence. Yoon's original study reported the instrument' construct validity and reliability for Korean nursing students; the explained variance for the factor analysis was 52.0%, and the instrument reliability using Cronbach alpha coefficient was .84. The YCTD was found to have strong reliability in several previous studies (Kim, 2012; Kim, 2010; Shin *et al.*, 2015; Yang *et al.*, 2009).

Data Analysis

Structural Equation Modelling (SEM) procedures in PLS was carried out to examine the validity and reliability of the YCTD. According to Wong (2016), PLS-SEM is primarily used to explain the variance in the dependent variables when examining the model. Therefore, considering YCTD based on the CCTDI which is a theory driven instrument, measurement model of PLS-SEM was used to test the YCTD's construct validity. Kyriazos (2018) and Fornell and Bookstein (1982) indicated that there are various advantages of using PLS-SEM. First, it does not suffer from indeterminacy problems like other causal modeling techniques. Second, it is a nonparametric technique which does not assume normality of the data. Third, it does not require a large sample size like other causal modeling techniques. Lastly, it can be used to estimate models that use both formative and reflective indicators. According to Hair Jr *et al.* (2021), results can be presented in PLS-SEM as an assessment of the reliability and validity of the measurement model. In this study, the measurement model in PLS-SEM is assessed in terms of the loadings of the items, average variance extracted (AVE) for each variable, Cronbach's alphas, composite reliabilities, inter-correlations among the variables, chi-square statistic and associated probability (p), the root mean square error of approximation (RMSEA) index, the standardized root mean squared residual (SRMR), the comparative fit index (CFI), the coefficient of determination, Akaike's information criterion (AIC), and the Bayesian information criterion (BIC).

Furthermore, Chua and Chua (2017) endorsed the use of PLS-SEM analysis to establish the validity and reliability of the variables in a model prior to further data analysis. Hence, in this study, the convergent validity and discriminant validity for all variables in the measurement model are examined first and followed by the composite reliability and Cronbach's alpha internal consistency reliability of the variables. When the loading of the items for each variable is greater than .50 and the average variance extracted (AVE) for the variable is greater than .50, the convergent validity is achieved (Hair et al., 2019). When the inter-correlations among the variables in the model is smaller than .90, discriminant validity is achieved (Byrne, 2013; Chua & Chua, 2017). In case of multi-collinearity inter-correlation, the coefficient should be ($r \geq .090$) between each of the variables (Barbara, 2016). Besides, in order to achieve reliability, both the values of composite reliability and Cronbach's alpha should be greater than .70 (Hair Jr et al., 2021).

Measurement invariance analysis was conducted to examine whether the YCTD instrument measures critical thinking disposition equally across gender. Measurement invariance analysis tests the hypothesis that the same measurement model applies across different groups (Meredith & Teresi, 2006). Therefore, configural invariance was conducted between the unconstrained and fully constrained models.

Ethical Consideration

Before beginning the research, permission is obtained from the instrument developer to use his respective instrument in this study. The instrument developer was communicated via e-mail. Subsequently, necessary formalities were taken to seek approvals from the administrative department of sampled universities and also from the Head of the Departments of respective universities to carry out the study in the selected universities. Subsequently, official letters of permission from university administration and HODs was shown to the teachers to guide and help to conduct the study in selected universities in Quetta district, Balochistan.

The next step was the distribution of the questionnaire (YCTD) to the PG students. Prior to giving the questionnaire (YCTD) to the participants, the objectives and scope of the research was explicitly communicated to them in compliance with research ethics. They were also informed that all information and responses will be kept private, and that no participants will be identified.

FINDINGS

CTD is a latent construct measured by YCTD through seven subscales which are (i) confidence; (ii) eager; (iii) fairness; (iv) objectivity; (v) prudence; (vi) skepticism; and (vii) systemacity. Validity and reliability of YCTD through PLS-SEM analysis is presented in the following subsections in detail.

Measurement Model for the YCTD

The assessment of measurement model is conducted to ensure the quality criteria of YCTD. The measurement model for YCTD and its seven subscales are shown in Figure 3 on the following page.

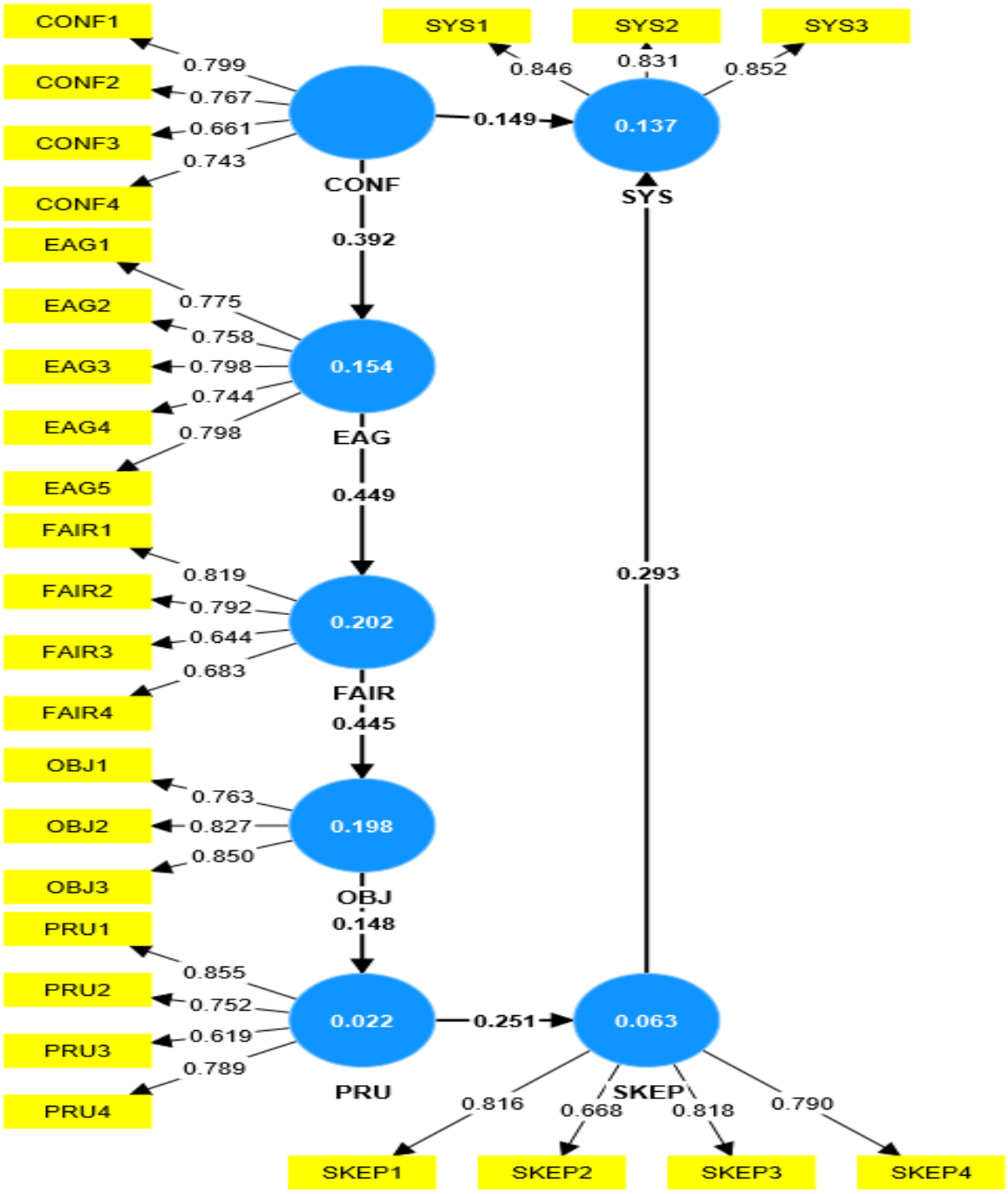

Figure 3. Measurement Model for the YCTD scale Through PLS-SEM

The assessment of the quality criteria starts with the evaluation of the factor loadings which is followed by establishing the construct reliability and validity. Factor loading refers to "the extent to which each of the items in the correlation matrix correlates with the given principal component. Factor loading can range from -1.0 to +1.0, with higher absolute values indicating a higher correlation of the items with the underlying factors" (Pett *et al.*, 2003, p.299). The YCTD had factor loading greater than the recommended value of >.50 (Hair Jr *et al.*, 2021). Hence, no items were further removed. Table 2 shows factor loading with values ranging from .644 to .864.

Table 2. Factor Loading for All Items of YCTD through PLS-SEM

Latent Variables	Factor Loadings
Confidence1	0.768
Confidence2	0.743
Confidence3	0.697
Confidence4	0.767
Eager 5	0.775
Eager 6	0.758
Eager 7	0.798
Eager 8	0.744
Eager 9	0.797
Fairness 10	0.819
Fairness 11	0.792
Fairness12	0.644
Fairness 13	0.683
Objectivity 14	0.765
Objectivity 15	0.830
Objectivity 16	0.847
Prudence 17	0.806
Prudence 18	0.776
Prudence 19	0.675
Prudence 20	0.814
Skepticism 21	0.815
Skepticism 22	0.671
Skepticism 23	0.818
Skepticism 24	0.790
Systemacity 25	0.847
Systemacity 26	0.817
Systemacity 27	0.864

Indicator Multicollinearity

Variance Inflation Factor (VIF) statistics is utilized to assess the multicollinearity in the indicators (Fornell & Bookstein, 1982). According to Hair Jr *et al.* (2021), multicollinearity is not a serious issue

if the value for VIF is below 5. Table 3 presents the VIF values for the indicators in the YCTD and shows that the multicollinearity statistics (VIF) outer values for YCTD are less than 5. Therefore, it reveals that the VIF for each of the subscales is below the recommended threshold and is at an acceptable level (see Table 3)

Table 3. Multicollinearity Statistics (VIF) Outer Values for YCTD Scale

Latent Variable	VIF
Confidence 1	1.485
Confidence 2	1.525
Confidence 3	1.302
Confidence 4	1.35
Eager 1	1.642
Eager 2	1.624
Eager 3	1.922
Eager 4	1.712
Eager 5	1.684
Fairness 1	1.405
Fairness 2	1.548
Fairness 3	1.346
Fairness 4	1.321
Objectivity 1	1.360
Objectivity 2	1.653
Objectivity 3	1.575
Prudence 1	1.387
Prudence 2	1.702
Prudence 3	1.548
Prudence 4	1.631
Skepticism 1	1.616
Skepticism 2	1.346
Skepticism 3	1.677
Skepticism 4	1.568
Systemacity 1	1.677
Systemacity 2	1.641
Systemacity 3	1.790

Convergent Validity

Convergent validity is the degree to which multiple attempts to measure the same concept are in agreement (Bagozzi *et al.*, 1991). When the Average Variance Extracted (AVE) value is greater than or equal to the recommended value of .50, items converge to measure the underlying construct and hence convergent validity is established (Fornell & Bookstein, 1982). Convergent validity results based on the AVE statistics in the current study shows that all the subscales of YCTD have AVE values greater than the recommended value of .50. (see Table 4). The value of AVE of Confidence is .554, Eager .600, Fair 0.545, Objectivity .664, Prudence .593, Skepticism .602 and Systemacity .711 which is greater than 0.50. Moreover, as part of the measurement model evaluation, no item was removed

from the analysis because all the items have acceptable factor loadings (>0.50). Therefore, the convergent validity for YCTD was achieved.

Table 4. Convergent Validity of All the Subscales of YCTD

Latent Variables	Average Variance Extracted (AVE)
Confidence	0.554
Eager	0.600
Fair	0.545
Objectivity	0.664
Prudence	0.593
Skepticism	0.602
Systemacity	0.711

Discriminant Validity

Discriminant validity is "the degree to which measures of different concepts are distinct". The notion is that "if two or more concepts are unique, than valid measures of each should not correlate too highly" (Bagozzi *et al.*, 1991, p. 425)

Fornell and Larker Criterion

Table 5 demonstrates the discriminant validity of all subscales of YCTD which are greater than its correlation with each of the other YCTD subscales. For Confidence, the square root value of AVE is .744, for Eager, the square root value of AVE is .775, for Fair .738, Objectivity is .815, Prudence .770, Skepticism .776 while for Systemacity is .843. Thus, it reveals strong support for the establishment of discriminant validity for the YCTD.

Table 5. Discriminant Validity of All Subscales of YCTD

		1	2	3	4	5	6	7
1	Confidence	***.744***						
2	Eager	.392	***.775***					
3	Fair	.323	.449	***.738***				
4	Objectivity	.342	.464	.445	***.815***			
5	Prudence	.259	.246	.240	.136	***.770***		
6	Skepticism	.330	.498	.414	.430	.238	***.776***	
7	Systemacity	.238	.308	.319	.349	.062	.342	***.843***

Note: Bold and Italic represents the square root of AVE

Inter-correlation and Cross Loadings of YCTD

Discriminant validity is also calculated through the examination of the inter-correlation among the subscales. According to Table 6, the inter-correlation coefficients among all the subscales of YCTD were less than .90 which ranged from 0.062 as lowest value and 0.449 as the highest value indicated that there is no significant multicollinearity problem. All correlation coefficient values are statistically significant. Table 7 shows cross loadings for all the items of YCTD from three subscales that are stronger than the underlying subscales to which they belong instead of the other subscales

in the instrument. Hence, based on the evaluation of cross loadings, discriminant validity is attained for the YCTD.

Table 6. Inter-correlation of YCTD

Latent Variables	1	2	3	4	5	6	7
1. Confidence							
2. Eager	0.392						
3. Fairness	0.323*	0.449					
4. Objectivity	0.342	0.464	0.445				
5. Prudence	0.259**	0.246**	0.240**	0.136**			
6. Skepticism	0.330	0.498	0.414	0.430	0.238**		
7. Systemacity	0.238**	0.308*	0.319*	0.349	0.062**	0.342	

Table 7. Cross Loadings of All the items of YCTD

	Confidence	Eager	Fairness	Objectivity	Prudence	Skepticism	Systemacity
Confidence1	0.768	0.34	0.264	0.288	0.159	0.27	0.23
Confidence 2	0.743	0.241	0.261	0.199	0.153	0.216	0.208
Confidence 3	0.697	0.222	0.148	0.254	0.228	0.182	0.126
Confidence 4	0.767	0.338	0.277	0.265	0.226	0.294	0.148
Eager 1	0.336	0.775	0.378	0.331	0.172	0.386	0.224
Eager 2	0.326	0.758	0.297	0.325	0.152	0.295	0.208
Eager 3	0.264	0.798	0.328	0.437	0.215	0.457	0.215
Eager 4	0.227	0.744	0.308	0.341	0.16	0.378	0.269
Eager 5	0.342	0.797	0.408	0.369	0.246	0.411	0.276
Fairness 1	0.315	0.482	0.819	0.403	0.182	0.417	0.296
Fairness 2	0.189	0.288	0.792	0.36	0.256	0.295	0.189
Fairness 3	0.155	0.222	0.644	0.184	0.104	0.197	0.171
Fairness 4	0.257	0.253	0.683	0.307	0.148	0.247	0.262
Objectivity 1	0.26	0.326	0.317	0.765	0.163	0.269	0.291
Objectivity 2	0.254	0.387	0.336	0.83	0.082	0.339	0.265
Objectivity 3	0.315	0.415	0.423	0.847	0.093	0.429	0.296
Prudence 1	0.235	0.226	0.223	0.186	0.806	0.259	0.088
Prudence 2	0.17	0.174	0.137	0.068	0.776	0.144	0.047
Prudence 3	0.135	0.149	0.097	0.032	0.675	0.061	-0.019
Prudence 4	0.221	0.19	0.229	0.079	0.814	0.193	0.036
Skepticism 1	0.321	0.385	0.33	0.366	0.221	0.815	0.286
Skepticism 2	0.187	0.304	0.215	0.207	0.114	0.671	0.212
Skepticism 3	0.279	0.458	0.39	0.354	0.175	0.818	0.298
Skepticism 4	0.22	0.384	0.327	0.379	0.212	0.79	0.257
Systemacity 1	0.171	0.222	0.244	0.278	0.041	0.327	0.847
Systemacity 2	0.234	0.278	0.245	0.268	0.056	0.256	0.817
Systemacity 3	0.203	0.284	0.316	0.334	0.06	0.279	0.864

Heterotrait-Monotrait Ratio (HTMT) and Standardized Root Mean Square Residual

The HTMT of all the items of YCTD is less than the required value (0.85). The HTMT ratio for Confidence is 0.482, Eager 0.530, Fair 0.566, Objectivity 0.157, Prudence 0.267, whereas Skepticism has the HTMT ratio of 0.429 which is > 0.85. However, the threshold of the HTMT has been debated in the existing literature; Kline (2011) suggested a threshold of .85 or less, while Teo *et al.* (2008) recommended a liberal threshold of 0.90 or less. The HTMT ratio for this study is less than recommended threshold of .90 (see Table 8).

To check the goodness of fit, the SRMR value was estimated. SRMR is used to calculate as the measure of Model fit. SRMR value of less than 0.10 or 0.08 is considered as the good fit (Byrne, 2013; Hu & Bentler, 1999). Table 9 shows the SRMR values of YCTD which is acceptable as per recommended level. The SRMR value for present study is 0.065 for saturated model and 0.133 for estimated model. Thus, the SRMR value for YCTD is showing a good fit.

Table 8. Heterotrait-Monotrait Ratio (HTMT) of All the Subscales of YCTD

		1	2	3	4	5	6	7
1	Confidence							
2	Eager	*0.482*						
3	Fairness	0.416	*0.530*					
4	Objectivity	0.452	0.586	*0.566*				
5	Prudence	0.322	0.291	0.285	*0.157*			
6	Skepticism	0.421	0.611	0.508	0.544	*0.267*		
7	Systemacity	0.315	0.379	0.404	0.45	0.088	*0.429*	

Table 9. Standardized Root Mean Square Residual (SRMR) of YCTD

	Saturated Model	Estimated Model
SRMR	0.065	0.133
d_ULS	1.594	6.685
d_G	0.437	0.558
Chi-Square	1032.104	1219.993
NFI	0.732	0.683

Reliability Analysis

The two most commonly used methods to establish reliability are the Cronbach Alpha and Composite reliability (CR). After assessing the convergent and discriminant validity of YCTD, the final step is to determine the reliability of the YCTD and its subscales. Reliability of the construct is achieved when both the values of composite reliability and Cronbach's alpha should be greater than .70 (Hair Jr *et al.*, 2021). Table 10 shows that the Cronbach Alpha and the Composite reliability (CR) for all the subscales of YCTD yields significant value higher than 0.7. The Cronbach Alpha ranged from .730 to .797 whereas the CR ranged from .826 to .882. Thus, both indicators of reliability (Cronbach Alpha, CR) have reliability statistics over the required threshold of 0.70 (Hair *et al.*, 2006). Hence, construct reliability is established.

Table 10. Construct Reliability Analysis

	Cronbach Alpha	Composite Reliability
Confidence	0.733	0.832
Eager	0.834	0.882
Fair	0.730	0.826
Objectivity	0.747	0.855
Prudence	0.782	0.853
Skepticism	0.779	0.857
Systemacity	0.797	0.880

Test of Measurement Invariance (MI)

The Tucker-Lewis Index (TLI) and Comparison Fit Index (CFI) are comparative indexes that compare the fit of the model under examination to the fit of a baseline model. If the CFI and TLI values are >.90, the fit is deemed adequate, and if they are >.95, the fit is considered better (Hu & Bentler, 1999; Van de Schoot *et al.*, 2012). The CFI is set to 1.0 if the $\chi 2 <$ df is true, making it a normed fit index (Van de Schoot *et al.*, 2012). Moreover, absolute indices, such as the RMSEA, are used to assess the closeness of fit. Those with RMSEA scores <.06 are regarded "good," while those with values between .08 are rated "mediocre" (Teo & Kam, 2014). There are indexes based on information theory, such as the Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC). BIC/AIC are applied to compare models and make trade-offs between model fit and complexity. A lower BIC/AIC number suggests a better fit/complexity trade-off (Teo & Kam, 2014). The chi-square test is the most often used test to verify global model fit, and it is also applied to compare different fixed models. However, it has also been found sensitive to vast sample size (Chen, 2007).

Table 11 provides the results of the measurement invariance (MI) analysis between Group 1 (unconstrained) and Group 2 (fully constrained) based on various fit indices. The chi-square values ($\chi 2$) for both groups were significant ($p < .001$), indicating that the null hypothesis of exact measurement invariance is rejected. This suggests that there are differences in the measurement model across the two groups (male/female).

Although the chi-square test is sensitive to sample size, other fit indices provide additional information. The comparative fit index (CFI) and Tucker-Lewis Index (TLI) values were above .87 for both groups, which suggests acceptable model fit. The root mean square error of approximation (RMSEA) values were close to .07 for both groups, indicating a reasonable fit. However, it should be noted that RMSEA values above .06 may suggest a mediocre fit.

The Akaike information criterion (AIC) was lower for Group 1 (1285.560) compared to Group 2 (1296.005), indicating a better fit for the unconstrained model. Overall, the results suggest that there are differences in the measurement model across the two groups, indicating a lack of measurement invariance. These findings highlight the need for cautious interpretation and comparison of scores between the male and female, as the construct may not be measured equivalently across gender (see Table 11).

Table 11. Summary of Measurement Invariance (MI)

	Group 1 (unconstrained)	Group 2 (fully constrained)
χ2	2.699	2.634
df	412	434
p	.000	.000
CFI	.886	.887
TLI	.875	.879
RMSEA	.073	.074
AIC	1285.560	1296.005

Note. χ2= chi-square; df = degree of freedom; CFI= Comparison Fit Index; TLI= Tucker-Lewis Index; RMSEA= Root Mean Square Error of Approximation AIC= Akaike Information Criterion; $p<0.05$

DISCUSSION

Comparing the previous efforts of Shin *et al.*, (2015); Yang *et al.*, (2009); Yoon, (2008) with this research, the authors validated the YCTD instrument (English version) with a sample of 390 PG university students in the context of Pakistan. Primarily, the present study provides evidence of the cross-sectional validity and construct validity of the seven-factor model of the YCTD in the context of Pakistan. It also provides its measurement invariance (factor loading, configural) across different subgroups of participants. These findings are consistent with previous research on the YCTD instrument, which has shown that it is a reliable and valid measure of critical thinking disposition in a variety of cultural contexts (Yoon, 2013; 2018). The hypothesis was confirmed in terms of the instrument's convergent validity.

One of the strengths of this study is that it provides evidence of the measurement model in PLS-SEM which shows that the seven subscales of the YCTD are valid indicators of critical thinking disposition in the Pakistani context. The measurement invariance analyses revealed that the factor structure, factor loadings, and intercepts of the YCTD were consistent across different subgroups of participants in Pakistan, indicating that the instrument is invariant and can be used to compare critical thinking disposition levels across different groups of participants. This is an important finding as it suggests that the YCTD instrument can be used to measure critical thinking disposition in diverse groups which is especially relevant in the Pakistani context where there are significant cultural and socioeconomic differences.

To begin with, the reliability coefficient (Cronbach's Alpha > 0.500) in this study indicated good reliability in all aspects. This is in line with previous research findings of Shin *et al.*, (2015); Yang *et al.*, (2009); Yoon, (2008) . The authors did not rely solely on Cronbach's alpha values in a series of efforts to test for validity. However, both findings are positive indicators that support the YCTD validity. The item composite reliability values (>0.70) indicated that each item is related to the measured subscale's total values. As a result, the authors conducted factor loadings which is commonly used in similar research in a variety of subjects. The significant differences in correlation coefficient between the seven subscales of YCTD suggest that the subscales are related to different aspects of the construct. The heterogeneity of these relationships can be interpreted as strong evidence for the variation into seven subscales, despite the lack of a clear and obvious pattern in these relationships.

Moreover, the results of the present study strengthened the logically based and empirically validated division of the seven subscales (objectivity, prudence, systematicity, intellectual eagerness,

intellectual fairness, healthy skepticism, CT self-confidence). Results also reveal that the Chi-square changes between unconstrained model (2.699) and fully constrained model (2.634) are not significant. Therefore, the measurement model is invariant by gender.

Moreover, present study was conducted with university students to investigate students' YCTD level because critical thinking disposition and critical thinking skill are closely correlated with academic success and affect achievement in the academic domain (Abrami *et al.*, 2015; Facione *et al.*, 2016; Pozhhan *et al.*, 2019).

The findings of this study have important implications for educators and policymakers in Pakistan, as they highlight the importance of promoting and measuring critical thinking in the education system. Previous research has shown that critical thinking is an important skill for success in academic and professional settings (Elder & Paul, 2001; Facione, 2011) and that interventions aimed at improving critical thinking can have positive effects on academic performance and success (Abrami *et al.*, 2008; Kuhn, 2015). The YCTD instrument provides a tool for evaluating the effectiveness of such interventions and can be used to monitor changes in critical thinking disposition over time.

As a result of these findings, we believe it would be important to investigate the evolution of CTD over time and through various stages of schooling to see if certain dimensions stay constant while others increase significantly as a student progress through the educational system. We assume that as compared to other available CTD instruments, YCTD is useful for guiding interventions at the individual level. As each student will be assessed on three distinct dimensions (objectivity, prudence, systematicity, intellectual eagerness, intellectual fairness, healthy skepticism, CT self-confidence), the strengths and weaknesses of each student can be explained and further addressed.

In summary, the findings of this study provide strong evidence of the validity and reliability of the YCTD instrument in the Pakistani context. The measurement invariance analyses also suggest that the YCTD instrument can be used to compare critical thinking disposition levels across different subgroups, making it a useful tool for measuring critical thinking disposition in diverse groups. These findings have implications for educators and policymakers in Pakistan, as they highlight the importance of promoting and measuring critical thinking in the education system and provide a tool for evaluating the effectiveness of interventions aimed at improving critical thinking skills.

IMPLICATIONS

The findings of this study have important practical implications for educators and policymakers in Pakistan. By validating the YCTD instrument and demonstrating its reliability and validity, this study provides a valuable tool for measuring and promoting critical thinking in the education system. Educators can utilize the YCTD to assess students' critical thinking disposition and identify areas for improvement. The instrument can guide interventions at the individual level by helping educators understand students' strengths and weaknesses across the seven subscales. Policymakers can consider incorporating the measurement of critical thinking disposition into educational policies and curriculum development, recognizing its significance for academic and professional success. Furthermore, the YCTD can be used to monitor changes in critical thinking disposition over time, allowing for the evaluation of interventions aimed at enhancing critical thinking skills.

The theoretical implications of this study contribute to the existing literature on critical thinking disposition. By validating the seven-factor model of the YCTD in the Pakistani context, this study strengthens the logical basis and empirical evidence supporting the division of critical thinking disposition into distinct subscales. The findings reinforce the notion that critical thinking disposition is multifaceted and comprises various dimensions such as objectivity, prudence, systematicity, intellectual eagerness, intellectual fairness, healthy skepticism, and CT self-confidence. This supports the conceptualization of critical thinking as a complex construct with different facets. Additionally, the measurement invariance analyses across different subgroups of participants underscore the robustness of the YCTD instrument, suggesting its applicability in diverse cultural and socioeconomic contexts. These theoretical insights contribute to a deeper understanding of critical thinking disposition and its measurement, enhancing the field of critical thinking research.

LIMITATIONS

Our study, like any other, has some limitations. The biggest limitation is that we were unable to evaluate students' CTD through any other means i.e., writing product, critical thinking test. YCTD is asked for self-assessment and it is obvious that many individuals lack proficiency in self-assessment (Ma & Winke, 2019), therefore, it is a limitation. Another limitation of this study is that the sample was not representative of the entire Pakistani population and was limited to participants who were recruited through the universities of Balochistan, Pakistan. Additionally, the study did not investigate the predictive validity of the YCTD, and future research should investigate how the YCTD is related to other measures of critical thinking as well as how it relates to academic performance and success.

SUGGESTIONS

Despite some limitations, YCTD may enable researchers to study individual differences because it is a brief, trustworthy, and valid CTD self-assessment measure. Firstly, further investigation is warranted to examine the longitudinal development of critical thinking disposition across different stages of schooling, tracking how specific dimensions may change or remain stable over time. This would provide valuable insights into the growth and stability of critical thinking disposition throughout an individual's educational journey. Further research can also use Rasch Measurement Model to ascertain the rating scale functionality, range of ability levels that can be measured, unidimensionality to test the data when analysed with Rasch Model to suggest a unidimensional structure. Additionally, exploring the relationship between critical thinking disposition and critical thinking skills could shed light on how these two constructs interact and influence each other. Moreover, extending the application of the YCTD instrument to other cultural contexts would enhance its cross-cultural validity and allow for comparative studies on critical thinking disposition across diverse populations. Lastly, investigating the effectiveness of interventions aimed at improving critical thinking disposition using the YCTD as an outcome measure would provide evidence-based guidance for educators and policymakers seeking to enhance critical thinking skills among students. Despite some limitations and further research avenues, it is believed that the current study will contribute to the validation studies of critical thinking in various aspects and settings especially in the context of Pakistan.

CONCLUSION

In conclusion, the findings of this study provide strong evidence of the validity and reliability of the YCTD instrument in the Pakistani context. The measurement invariance analyses also suggest that the YCTD instrument can be used to compare critical thinking disposition levels across different subgroups, making it a useful tool for measuring critical thinking disposition in diverse groups. The findings of the study indicated that the YCTD instrument is a valid and reliable measure of critical thinking disposition in the Pakistani context, and it demonstrated measurement invariance across different subgroups of participants. These findings have important implications for promoting and measuring critical thinking in the education system in Pakistan, and for evaluating the effectiveness of interventions aimed at improving critical thinking skills. Moreover, by promoting and measuring critical thinking in the education system, educators and policymakers can help prepare students for academic and professional success in today's complex and rapidly changing world.

Overall, the results of this study support the use of the YCTD instrument as a valuable tool for measuring critical thinking disposition in the Pakistani context and suggest that it can be used to compare critical thinking disposition levels across different groups of participants. However, future research should explore the predictive validity of the YCTD instrument and its relationship with other measures of critical thinking, as well as its relationship with academic performance and success.

REFERENCES

Abrami, P. C., Bernard, R. M., Borokhovski, E., Waddington, D. I., Wade, C. A., & Persson, T. (2015). Strategies for teaching students to think critically: A meta-analysis. *Review of educational research*, *85*(2), 275-314.

Abrami, P. C., Bernard, R. M., Borokhovski, E., Wade, A., Surkes, M. A., Tamim, R., & Zhang, D. (2008). Instructional interventions affecting critical thinking skills and dispositions: A stage 1 meta-analysis. *Review of educational research*, *78*(4), 1102-1134.

Babbie, E. R. (2020). *The practice of social research*. Cengage learning.

Bagozzi, R. P., Yi, Y., & Phillips, L. W. (1991, p. 425). Assessing construct validity in organizational research. *Administrative science quarterly*, 421-458.

Bali, M. (2015). Critical thinking through a multicultural lens: Cultural challenges of teaching critical thinking. In *The Palgrave handbook of critical thinking in higher education* (pp. 317-334). Springer.

Barta, A., Fodor, L. A., Tamas, B., & Szamoskozi, I. (2022). The development of students critical thinking abilities and dispositions through the concept mapping learning method – A meta-analysis. *Educational Research Review*, *37*, 100481. https://doi.org/https://doi.org/10.1016/j.edurev.2022.100481

Bravo, M. J., Galiana, L., Rodrigo, M. F., Navarro-Pérez, J. J., & Oliver, A. (2020). An adaptation of the Critical Thinking Disposition Scale in Spanish youth. *Thinking Skills and Creativity*, *38*, 100748. https://doi.org/https://doi.org/10.1016/j.tsc.2020.100748

Butler, H. A., Pentoney, C., & Bong, M. P. (2017). Predicting real-world outcomes: Critical thinking ability is a better predictor of life decisions than intelligence. *Thinking Skills and Creativity*, *25*, 38-46.

Byrne, B. M. (2013). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming*. Psychology Press.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural equation modeling: a multidisciplinary journal*, *14*(3), 464-504.

Chua, Y. P., & Chua, Y. P. (2017). Developing a grounded model for educational technology leadership practices. *Egitim ve Bilim*, *42*(189).

Du, J., Zhu, X., Wang, J., Zheng, J., Zhang, X., Wang, Z., & Li, K. (2022). History-taking level and its influencing factors among nursing undergraduates based on the virtual standardized patient testing results: Cross sectional study. *Nurse Education Today*, *111*, 105312. https://doi.org/https://doi.org/10.1016/j.nedt.2022.105312

Elder, L., & Paul, R. (2001). Critical thinking: Thinking to some purpose. *Journal of Developmental Education*, *25*(1), 40.

Ennis, R. H., Millman, J., & Tomko, T. N. (2005). *Cornell Critical Thinking Tests Levels X & Level Z Manual*. Critical Thinking Company.

Facionc, P., & Facione, N. (1992). The California Critcal Thinking Dispositons Inventory (C7D1). *Milbrae, CA.: The California Academic Pres*.

Facione, P. (1990). Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction (The Delphi Report).

Facione, P. (2000). The Disposition Toward Critical Thinking: Its Character, Measurement, and Relationship to Critical Thinking Skill lnformal Logic Vol. 20.

Facione, P., Gittens, C., Facione, P., & Gittens, C. (2016). The power of critical thinking. *Facione PA, Gittens CA. Think critically. Boston: Pearson Education*, 1-17.

Facione, P. A. (2011). Critical thinking: What it is and why it counts. *Insight assessment*, *2007*(1), 1-23.

Fornell, C., & Bookstein, F. L. (1982). Two structural equation models: LISREL and PLS applied to consumer exit-voice theory. *Journal of Marketing research*, *19*(4), 440-452.

Garg, R. (2016). Methodology for research I. *Indian journal of anaesthesia*, *60*(9), 640-645. https://doi.org/10.4103/0019-5049.190619

[Record #1249 is using a reference type undefined in this output style.]

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European business review*, *31*(1), 2-24.

Hair Jr, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2021). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Sage publications.

Halpern, D. F. (2010). Halpern critical thinking assessment. *SCHUHFRIED (Vienna Test System): Moedling, Austria. Http://www. Schuhfried. com/vienna-Test-System-vts/all-Tests-fromaz/test/hcta-Halpern-Critical-Thinking-Assessment*, *1*.

Hu, L. t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, *6*(1), 1-55.

İskifoğlu, G., & Ağazade, A. S. (2013). Translation and validation of a Turkish version of the California Critical Thinking Disposition Inventory. *Social Behavior and Personality: an international journal*, *41*(2), 187-196.

Kim, D. (2012). Improvement in problem solving and critical thinking among Korean nursing students over an academic year. *Educational Research Journal*, *2*(8), 257-265.

Kim, M. Y., & Byun, E. K. (2019). Influence of academic self-efficacy, critical thinking disposition, and learning motivation on problem solving ability in nursing students. *Journal of the Korea Academia-Industrial cooperation Society*, *20*(1), 376-383.

Kim, S. A., Hong, E., Kang, G. Y., Brandt, C., & Kim, Y. (2020). Effect of Korean nursing students' experience of incivility in clinical settings on critical thinking. *Heliyon*, *6*(7), e04367. https://doi.org/https://doi.org/10.1016/j.heliyon.2020.e04367

Kim, Y.-M. (2010). Factors influencing problem solving abilities of freshmen nursing students. *Journal of Korean Academy of Nursing Administration*, *16*(2), 190-197.

Kline, R. B. (2011). Principles and practice of structural equation modeling (3. Baskı). *New York, NY: Guilford*, *14*, 1497-1513.

Kuhn, D. (2015). Thinking together and alone. *Educational researcher*, *44*(1), 46-53.

Kyriazos, T. A. (2018). Applied psychometrics: sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general. *Psychology*, *9*(08), 2207.

Li, Y., Li, K., Wei, W., Dong, J., Wang, C., Fu, Y., Li, J., & Peng, X. (2021). Critical thinking, emotional intelligence and conflict management styles of medical students: A cross-sectional study. *Thinking Skills and Creativity*, *40*, 100799. https://doi.org/https://doi.org/10.1016/j.tsc.2021.100799

Ma, W., & Winke, P. (2019). Self-assessment: How reliable is it in assessing oral proficiency over time? *Foreign Language Annals*, *52*(1), 66-86. https://doi.org/https://doi.org/10.1111/flan.12379

Meredith, W., & Teresi, J. A. (2006). An essay on measurement and factorial invariance. *Medical care*, S69-S77.

Nair, G. G., & Stamler, L. L. (2013). A conceptual framework for developing a critical thinking self-assessment scale. *Journal of Nursing Education*, *52*(3), 131-138.

Norris, S. P. (2003). The meaning of critical thinking test performance: The effects of abilities and dispositions on scores. *Critical thinking and reasoning: Current research, theory and practice*, 315-329.

Papp, K. K., Huang, G. C., Clabo, L. M. L., Delva, D., Fischer, M., Konopasek, L., Schwartzstein, R. M., & Gusic, M. (2014). Milestones of critical thinking: a developmental model for medicine and nursing. *Academic Medicine*, *89*(5), 715-720.

Paul, R. W., & Binker, A. (1990). *Critical thinking: What every person needs to survive in a rapidly changing world*. ERIC.

Pett, M. A., Lackey, N. R., & Sullivan, J. J. (2003, p.299). *Making sense of factor analysis: The use of factor analysis for instrument development in health care research*. sage.

Pincus, K. V., Stout, D. E., Sorensen, J. E., Stocks, K. D., & Lawson, R. A. (2017). Forces for change in higher education and implications for the accounting academy. *Journal of Accounting Education*, *40*, 1-18.

Poondej, C., & Lerdpornkulrat, T. (2015). The Reliability and Construct Validity of the Critical Thinking Disposition Scale.

Pozhhan, P., Goodarzi, K., & Roozbehani, M. (2019). Investigating the relationship between critical thinking and academic achievement in male students of the Islamic Azad university of Ahvaz branch. *International Journal of Philosophy and Social-Psychological Sciences*, *5*(3), 9-17.

Qiuhuan, Y. (2008). The relationship among critical thinking, communication ability and perceived self-efficacy in nursing students. https://doi.org/https://kns.cnki.net/kns/brief/Default_Result.aspx?code=SCDB

Rebele, J. E., & Pierre, E. K. S. (2019). A commentary on learning objectives for accounting education programs: The importance of soft skills and technical knowledge. *Journal of Accounting Education*, *48*, 71-79.

Ren, X., Tong, Y., Peng, P., & Wang, T. (2020). Critical thinking predicts academic performance beyond general cognitive ability: Evidence from adults and children. *Intelligence*, *82*, 101487.

Shin, H., Park, C. G., & Kim, H. (2015). Validation of Yoon's critical thinking disposition instrument. *Asian nursing research*, *9*(4), 342-348.

Shin, K., Jung, D. Y., Shin, S., & Kim, M. S. (2006). Critical thinking dispositions and skills of senior nursing students in associate, baccalaureate, and RN-to-BSN programs. *Journal of Nursing Education*, *45*(6), 233.

Siegel, H. (1997). *Rationality redeemed?: Further dialogues on an educational ideal*. Psychology Press.

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/https://doi.org/10.1016/j.tsc.2012.09.002

Tafazzoli, M., Fakari, F. R., Ramazanzadeh, M., & Sarli, A. (2015). The relationship between critical thinking dispositions and academic achievement in Iranian midwifery students. *Nursing Practice Today*, *2*(3), 88-93.

Teo, T., & Kam, C. (2014). A measurement invariance analysis of the general self-efficacy scale on two different cultures. *Journal of Psychoeducational Assessment*, *32*(8), 762-767.

Teo, T. S., Srivastava, S. C., & Jiang, L. (2008). Trust and electronic government success: An empirical study. *Journal of management information systems*, *25*(3), 99-132.

Thomas, C. L., & Hayes, A. R. (2021). Using Exploratory Structural Equation Modeling to Investigate the Construct Validity of the Critical Thinking Disposition Scale. *Journal of Psychoeducational Assessment*, *39*(5), 640-648. https://doi.org/10.1177/0734282921990564

Van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European journal of developmental psychology*, *9*(4), 486-492.

Walsh, C. M., & Hardy, R. C. (1997). Factor Structure Stability of the California Critical Thinking Disposition Inventory across Sex and Various Students' Majors. *Perceptual and Motor Skills*, *85*(3_suppl), 1211-1228. https://doi.org/10.2466/pms.1997.85.3f.1211

Walsh, C. M., Seldomridge, L. A., & Badros, K. K. (2007). California Critical Thinking Disposition Inventory: Further Factor Analytic Examination. *Perceptual and Motor Skills*, *104*(1), 141-151. https://doi.org/10.2466/pms.104.1.141-151

Werner, P. H. (1991). The Ennis-Weir critical thinking essay test: An instrument for testing and teaching. *Journal of Adolescent & Adult Literacy*, *34*(6), 494.

Wolcott, S. K., & Sargent, M. J. (2021). Critical thinking in accounting education: Status and call to action. *Journal of Accounting Education*, *56*, 100731. https://doi.org/https://doi.org/10.1016/j.jaccedu.2021.100731

Wong, K. K.-K. (2016). Mediation analysis, categorical moderation analysis, and higher-order constructs modeling in Partial Least Squares Structural Equation Modeling (PLS-SEM): A B2B Example using SmartPLS. *Marketing Bulletin*, *26*(1), 1-22.

Yang, S.-H., Lee, O.-C., Lee, W.-S., Yoon, J., Park, C.-S., & Lee, S.-J. (2009). Critical disposition and clinical competency in 3 nursing colleges with different education methods. *The Journal of Korean Academic Society of Nursing Education*, *15*(2), 149-158.

Yang, S. A. (2010). Critical thinking disposition and problem solving ability in nursing students [Article]. *Journal of Korean Academy of Nursing Administration*, *16*(4), 389-398. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84925688633&partnerID=40&md5=7b349144a932cc9ae3fec2eb433796c2

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*(2), 123-132. https://doi.org/https://doi.org/10.1016/S0020-7489(01)00019-0

Yockey, R. D. (2016). Validation Study of the Critical Thinking Dispositions Scale: A Brief Report. *North American Journal of Psychology*, *18*, 101.

Yoon, J. (2008). Development of an Instrument for the Measurement of Critical Thinking Disposition in Nursing. *The Catholic University of Education, Seoul (Korea)*, *69*.

Yoon, J., & Kim, K. H. (2013). Testing the structure of critical thinking abilities for korean university students: A validity study of the critical thinking disposition inventory. *Educational and psychological measurement, 73(1), 38-59*.

Yoon, J., Kim, K. H., & Kim, D. (2018). Cross-cultural validation of the critical thinking disposition inventory in korea and the USA. *International journal of testing, 18(4), 348-367*.