Journal of Professional & Applied Psychology *Original Article*

Urdu Translation and Adaptation of Cornell Critical Thinking Test-Level Z in Pakistan

Abdul Wahab Liaquat[1], Najam ul Hasan Abbasi[2], Siara Ferdous[3]

Abstract

This study aims to translate and adapt the Cornell Critical Thinking Test- Level Z (CCTT-Z) into Urdu language in Pakistani population. Phase-I of the study dealt with Urdu translation, back translation, and cross-language validation. Recommended protocols were followed for forward and backward translations. Results of cross-language validation based on a sample of college students (n=42, Mage=18.69 years) showed that scores on both Urdu and English language versions were positively correlated (r =.31, p<.05). Moreover, 45.2% of students reported the Urdu language version easier than the English version. Certain lingual changes were made to make the test more culturally compatible without disturbing its core structure. For the Phase-II, a sample of 367 students of undergraduate and graduate programs (Mage= 21.56 years) was obtained from colleges, universities, and a teaching hospital. The results showed that the test had a high difficulty (p=.38). Twelve items had non-significant point-biserial correlation coefficients and dropping them improved overall Cronbach alpha reliability of the test (from α=.47 to α=.59). Five-factor theoretical model and four-factor model (induction items removed) were examined through CFA. The test showed poor model fit and did not appear to have any meaningful factor structure. This finding was not unique as several other validation studies across various cultures also do not show theoretically predicted factor structure. The findings are discussed by examining the available literature on the cross-cultural validity studies done in several other countries.

Keywords: Cornell-Critical Thinking Test, Critical Thinking, Urdu Translation, Pakistan

Received: 07 July 2022; Revised Received: 24 August 2022; Accepted: 21 September 2022

[1]Assistant Professor, Department of Psychology, Government Gordon Graduate College, Rawalpindi, Pakistan.
[2]Associate Professor, Department of Academic Sciences, Mianyang Normal University, Sichuan, China.
[3]Lecturer, Department of Applied Psychology, National University of Modern Languages, Islamabad, Pakistan.

Corresponding Author Email:
wahabliaqat@gmail.com

Introduction

The impact of critical thinking is well established in many areas of life especially in academic spheres (D'Alessio et al., 2019; Haseli & Rezaii, 2013; Ren et al., 2020; Shahzadi et al., 2020; Shirazi & Heidari, 2019; Taghva et al., 2014). Various opinions exist about its nature, measurement, and range of application (Ennis, 1962, 1989, 2011a; Facione, 1990, 2011; McPeck, 1981; Sanders & Moulenbelt, 2011). For Facione (1990), critical thinking is "purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgment is based." There are some academics who dislike the idea that critical thinking comprises a set of general skills that can be taught (Anderson, 2015; McPeck, 1981), while there are others who strongly believe that training of critical thinking in academia and other areas of life is vital (Ennis, 1962, 1989, 2002, 2011a; Facione, 2011; Halpern, 2013).

Unfortunately, in various educational policies of Pakistan, the emphasis on fostering critical thinking has remained a rhetoric without any successful policy implementations. Practically, the Pakistani educational system still favors rote-learning where the art of questioning and systematic doubt is not encouraged (Akhtar, 2019). This also reflects in nation's higher education system where, in last 10 years, despite a dramatic increase in quantity of research publications, the quality of most of the research work with respect to originality of ideas and rigor of methods is still deficient (Nauman, 2017).

Research in critical thinking has resulted in development of many measurement instruments some of which are widely used to this day (Ennis & Millman, 1971; Facione et al., 1990; Halpern, 2010; Watson & Glaser, 2007).

The Cornell Critical Thinking Test (CCTT) by Ennis et al. (2005) is particularly relevant in the context of this study. This test is founded on the Cornell-Illinois model (Ennis, 1962, 2002, 2011b, 2015b; Ennis & Millman, 1971) that conceptualizes critical thinking as "reasonable and reflective thinking focused on deciding what to believe or do" (Ennis, 2015a, p. 32). The CCTT-levels X and Z were first developed in 1971 (Ennis & Millman, 1971). A revision came in 1985 (Ennis et al., 1985), and currently its 5th edition is being published (Ennis, et al., 2005). Level X is intended for students from 4th grade to grade 14, whereas level Z is more suitable for advanced high school students, undergraduate, and graduate students. The CCTT-Z is a 52-item 'general-content' critical thinking test that claims to measure five dimensions including: i) *deduction,* ii) *induction,* iii) *meaning and fallacies,* iv) *observation and credibility of sources,* and v) *definition and assumption identification.* The test manual reports number of evidence for its construct and criterion validity. The test has correlations that range from r = .25 to .79 with seven other critical thinking tests and correlations from r = .24 to .71 with aptitude, intelligence and admission tests (Ennis et al., 2005, p. 33). Several studies have tried to establish theoretically compatible factor structure of CCTT-Z but the five dimensions discussed by the authors of the test do not find an empirical support (Follman et al., 1970; Frisby, 1992; Leach et al., 2020; Michael et al., 1980; Verburgh, et al., 2013). So far, the Cornell test has been translated into various languages including Dutch (Verburgh et al., 2013), Japanese (Hirayama et al., 2010), and Turkish (Sahin et al., 2015). This study aims to translate the test in Pakistani population as no indigenously developed test is yet traceable.

In Pakistani institutes, mostly the admissions in MPhil and Ph.D. level programs are secured through appearing in admission tests conducted by universities or by a government approved private body called National Testing Service (NTS). The tests at the level of the NTS assess students' logical reasoning, and mathematical and analytical abilities. No independent instruments of critical thinking exist for public and for research purposes. Urdu, which is a national language of Pakistan, and currently, no Urdu language version of CCTT exists. There is scarcity of research in Pakistan specifically targeting critical thinking abilities. Mostly, the studies in critical thinking are restricted to foreign language learning and teaching (Rashid & Qaisar, 2017). Lack of availability of indigenous instrument provides a great hinderance for rigorous research in this field. The present study tries to fill this gap by translating a critical thinking instrument to make the construct more accessible to the local population for further research. In order to reach that, following objectives were set for this study.

1. To translate and adapt Cornell Critical Thinking Test (CCTT)-Level Z into Urdu language.
2. To examine the item difficulty and item discrimination indices of the Cornell Critical Thinking Test (CCTT)-Level Z (Urdu version).

3. To examine the factor structure of the Cornell Critical Thinking Test (CCTT)-Level Z (Urdu version).
4. To examine the reliability of the Cornell Critical Thinking Test (CCTT)-Level Z (Urdu version).

Method

The study comprised two phases. The first phase included translation of the test, pre-tryout, and a cross-language validation study. The second phase included determining of the psychometric properties of the translated version. For the purpose of translation of the test into Urdu language, recommended translation guidelines were adopted (Brislin, 1970, 1986). The first phase is detailed below.

Phase I

Translation and Adaptation

The purpose of the first phase was to achieve an equivalent Urdu language translation of the original English language version of the test. Another purpose was to modify anything culturally dissimilar—that could potentially affect performance on the test—in such a way as to make the Urdu version more culturally compatible, without disturbing the core structure of the test. The translation and adaptation of CCTT-Z was accomplished through the following steps.

Forward Translations

Three separate Urdu language translations of the test were done by three different bilingual experts. The experts included a lecturer from a university in the field of Psychology, an assistant professor of Psychology from a government college specializing in Educational Psychology, and an Urdu language lecturer from a government college with an experience of translating various literary works.

Review and Selection of an Appropriate Translation

Another panel comprising a university lecturer in psychology, and two PhD assistant professors of psychology reviewed all three translations that resulted in a merger of single translation.

Back Translation

An assistant professor in English language and literature from a government college back translated the final Urdu translated version into English. The two versions were then assessed by a PhD assistant professor of Psychology to determine their equivalence. Few changes were incorporated in the Urdu language version keeping in view the English translation.

Pre-tryout

A pre-tryout study was carried out on a small sample of college students (n = 15). Overall, students reported the test rather difficult that required quite a bit of effort. The recommended test administration time limit of 50 minutes was not sufficient to complete the test. Participants, on average took between 60 to 90 minutes to complete the test. Some language difficulties were observed regarding understanding of difficult words and terms, and slight changes were made in the test accordingly.

Selecting Final Translated Version of the Test

Considering the *emic* and *etic* sensitivity (Barnouw, 1982; Berry, 1969; Phillips & Luna, 1996), and to make it more culturally familiar, few changes were done in the test. These changes mainly concerned with replacing English names with indigenous names without affecting the core structure of the test. It was ensured that name replacements should be culturally value-neutral. Though common place Pakistani Muslim names were used (Islam is the majority religion in Pakistan) but all such names were avoided that might indicate religious, sectarian, or political attachments. The major changes brought in the Urdu version of the test are given in Table 1.

Table1

Major changes in the Urdu language adaptation of CCTT-Level Z

Serial No.	Type of Change	English Language Version	Urdu Language Version
		Section IA and IB	**Section IA and IB**
1	Name of person	Mr Pinder	Parvaiz Sahab/پرویز صاحب
2	Name of person	Mr Wilstings	Wajahat Sahab/وجاہت صاحب
		Section II	**Section II**
3	Name of person	Dobert	Daniyal/دانیال
4	Name of person	Algan	Ahsan/احسان
5	Name of city	Galton city	Fateh Abad/ فتح آباد
		Section III, IV, V	**Section III, IV, V**
6	Name of person	Dr E. E. Brown	Dr. Baqir/ڈاکٹر باقر
7	Name of person	Dr M. R. Kolter	Dr Kamran/ڈاکٹر کامران
8	Name of duckling	Mallard	Green headed duck/ سبز سر والی بطخ
9	Name of duckling	Pintail	Pin tailed duck/ نوکیلی دم والی بطخ
10	Name of duckling	Canvasback	White Back duck/ سفید پشت والی بطخ
		Section VI	**Section VI**
11	Name of person	Bill	Naeem/نعیم
12	Name of person	Joan	Junaid/جنید
13	Name of person	Mary	Sara/سارہ
14	Name of person	Jim	Nadeem/ندیم

Proper nouns for ducklings (i.e., Canvasback, Mallard, and Pintail) were mentioned in the English version of the test. It was rather difficult to translate those names into Urdu as equivalent names do not exist for such duckling types in Urdu language. As these names depict outward features of the ducks (for example, Canvasback has a white back like a canvas, and Pintail has a pointed tail), so the Urdu names that matched those characteristic features were used in the Urdu version. To make things more comprehensible, colored pictures of ducklings were also placed under their names in the Urdu version (Figure 1).

خوراک بمعہ کیڑے			مقرر کردہ نارمل خوراک			ہر گروہ میں بطخ کے چوزوں کی اصل تعداد	بطخوں کی اقسام
مردہ	بیمار	صحت مند	مردہ	بیمار	صحت مند		
2	2			1	3	8	سبز سر والی بطخ
3					3	6	
3			1		2	6	نوکیلی دم والی بطخ
3		1		1	3	8	
3	1				4	8	سفید پشت والی بطخ
3	1			1	3	8	
17	4	1	1	3	18	44	کل تعداد

Figure 1

Urdu Equivalent Names of Ducklings along with Colored Pictures

Cross-language Validation

A sample of students (n = 42) of BS program majoring in Physics from a government college was selected for the tryout and cross-validation study. The sample included 34 males and 8 females with a mean age of 18.69 years. The CCTT-Z (Urdu) was administered with standard instructions and no strict time limits were imposed. With a gap of one day, CCTT-Z (English) was administered on the same sample. The average test completion time for the English version was 41.09 minutes, while the average completion time for the Urdu version was 56 minutes. As the English version was administered after the Urdu version, familiarity with the test content while solving the English version cannot be ignored. The participants were asked about the difficulty level of the English version in comparison to Urdu language version on a 3-point scale including 1 = easier than Urdu, 2 = same as Urdu, and 3 = difficult than Urdu. 17 participants (40.5%) reported English version easier, 6 participants (14.3%) reported it as equal, and 19 participants (45.2%) reported the English language version difficult than the Urdu version.

Table 2

Descriptive Statistics for English and Urdu Versions of CCTT-Z (n = 42)

	Mean	SD	Skewness	Kurtosis
CCTT-Z (English)	20.9	3.0	.36	-.11
CCTT-Z (Urdu)	19.7	3.4	-.03	-.26

Note = Test scores calculated by using rights-only method of scoring. Each correct response was scored as 1 and each incorrect response was marked as 0.

Table 3

Phi-Correlation Coefficients (ϕ) between Each Item of CCTT-Z (English) and CCTT-Z (Urdu) (n = 42)

Item No.	(ϕ)	Item No.	(ϕ)	Item No.	(ϕ)	Item No.	(ϕ)	Item No.	(ϕ)	Item No.	(ϕ)
1	.21	10	.32*	19	.08	28	.24	37	.36*	46	.41*
2	.35*	11	.26	20	.27	29	.02	38	.48*	47	.09
3	.69*	12	.28	21	.23	30	-.01	39	.18	48	.31*
4	.05	13	-.02	22	.14	31	.31*	40	.02	49	-.31*
5	.19	14	.01	23	-.13	32	.36*	41	-.13	50	.26
6	.25	15	-.14	24	-.07	33	.29	42	.38*	51	.06
7	-.02	16	.05	25	.13	34	.49*	43	.13	52	.03
8	.07	17	.18	26	.46*	35	-.02	44	.23		
9	.16	18	-.08	27	.17	36	.07	45	.11		

**p<.05*

In Table 3, items 2, 3, 10, 26, 31, 32, 34, 37, 38, 42, 46, and 48 show significant positive correlation ($p<.05$). 10 items show negative correlations, while remaining all items show positive though non-significant correlations. There is also a significant positive correlation) Pearson *r*) between total scores of both versions of the test ($r = .31$, $p<.05$).

Phase-II

Psychometric Properties of CCTT-Z (Urdu)

Sample

The test was administered on 400 university and college students through convenience sampling method. The incomplete test forms were dropped (i.e., forms of participants who withdrew from study or those forms that were left blank). The suitable sample for data analysis consisted of 367 participants (males 36.5 %, females 63.5%), with age ranges from 18 to 33 years (*M*age = 21.56 years). The students belonged to BS, Masters, MPhil, and PhD programs and were studying in disciplines of Mass Communication, Mathematics, Physiotherapy and Psychology. Convenience sampling was used and data were collected from seven universities, two colleges, and a teaching hospital. The institutes were located in Islamabad, and the provinces of Punjab, Khyber Pakhtunkhwa, and Baluchistan.

Procedure

The APA ethical guidelines were kept in view throughout the whole planning of the study as well during collection of data. The study was approved by the Board of Advanced Studies and Research (BASR) of the International Islamic University, Islamabad**.** Formal permissions were obtained from selected universities and colleges for data collection. Written consent forms were distributed to research participants prior to administration of the test. Throughout different administration sessions, standard instructions were used as mentioned in the administration manual of the test (Ennis et al., 2005). It was ensured that identities of the participants would not be revealed, information provided by them would be confidential, and there were no commercial or other hidden aspects of the research. Purpose of the study was delineated to the participants and their queries were addressed. Participants were also provided with researchers' emails to address any future queries regarding this study. The tests were administered in classrooms. The test manual recommends a 50-minute time limit for the completion. Though strict time restrictions were imposed, the participants were encouraged about this recommendation. In fact, in most of the cases, the average completion time exceeded one hour. The institutional

administrations and the participants were thanked for their cooperation and effort.

Results

Table 4

Means and Standard Deviations of CCTT-Z (Urdu) Items (N =367)

Item No.	M	SD	Item No.	M	SD	Item No.	M	SD	Item No.	M	SD
1	.37	.48	14	.28	.45	27	.40	.49	40	.33	.47
2	.65	.48	15	.52	.50	28	.34	.47	41	.42	.49
3	.14	.34	16	.33	.47	29	.41	.49	42	.47	.50
4	.51	.50	17	.59	.49	30	.28	.45	43	.31	.46
5	.43	.49	18	.23	.42	31	.33	.47	44	.36	.48
6	.25	.43	19	.32	.46	32	.33	.47	45	.36	.48
7	.34	.47	20	.29	.46	33	.53	.50	46	.45	.50
8	.59	.49	21	.17	.37	34	.40	.49	47	.47	.50
9	.28	.45	22	.26	.44	35	.31	.46	48	.39	.49
10	.62	.49	23	.36	.48	36	.51	.50	49	.42	.49
11	.36	.48	24	.40	.49	37	.36	.48	50	.44	.50
12	.32	.47	25	.39	.49	38	.43	.50	51	.29	.45
13	.32	.47	26	.49	.50	39	.28	.45	52	.29	.45

Table 4 shows that most of the mean scores range between .3 to .5 indicating high difficulty of the test.

Table 5

Point-Biserial Correlation (r_{bis}) and Item-Difficulty Information (p) for the Items of CCTT-Z (Urdu) (N = 367)

Item No.	r_{bis}	p	Item No.	r_{bis}	p	Item No.	r_{bis}	p	Item No.	r_{bis}	p
1	.28**	.37	14	.21**	.28	27	.18**	.42	40	.09	.33
2	.18**	.64	15	.21**	.52	28	.27**	.36	41	-.01	.42
3	.13*	.14	16	.17**	.33	29	.32**	.41	42	.23**	.46
4	.29**	.51	17	.19**	.19	30	.23**	.29	43	.31**	.31
5	.27*	.43	18	.09	.22	31	.03	.34	44	.15**	.35
6	.23*	.25	19	.26**	.27	32	.18**	.32	45	.23**	.35
7	.19**	.34	20	.10	.30	33	.35**	.53	46	.32**	.45
8	.23**	.59	21	.07	.16	34	.25**	.42	47	.28**	.46
9	.14*	.28	22	.22	.24	35	.08	.32	48	.19**	.38
10	.17**	.62	23	.17**	.35	36	.13*	.51	49	.28**	.45
11	.20**	.36	24	.08	.41	37	.06	.34	50	.20**	.43
12	.19**	.32	25	.23**	.38	38	.18**	.41	51	.28**	.41
13	.28**	.32	26	.03	.48	39	-.01	.29	52	.11*	.29

** $p<.01$; * $p<.05$

In Table 5, the pattern of correlation coefficients shows a heterogenous nature of the construct. The mean item difficulty for the test is $p=.38$. Taking into account the conventional standards (Kaplan & Saccuzzo, 2005, pp. 168-169; Riaz, 2017, p. 273), critical thinking in the present study appears to be a rather difficult construct. The Cronbach alpha reliability for full 52 items is $\alpha=.47$ that improves to $\alpha=.59$ if Induction subtest items are removed.

Table 6

Comparison between Model Fit for Four-Factor and Five-Factor Structures of CCTT-Z (Urdu) (N = 367)

Model	Items	χ 2	df	SRMR	TLI	CFI	RMSEA	LCI	UCI
Five Factor Model (default)	52	1893.79*	1264	.10	.49	.52	.037	.033	.040
Four Factor Model (M1)	35	719.39*	458	.10	.52	.56	.039	.034	.045

**p<.05*

Note: M1 = Four factor model is without induction subtest

Table 6 shows model fit comparison with Diagonal Weighted Least Square (DWLS) estimation method using R version 3.6.1 (R Core Team, 2019) to see factor loadings. The default model comprised five factors on which the test was actually based. The Chi-square statistic for both models is significant, while the other recommended model fit indices including SRMR, CFI, and TLI for both models show inappropriate fit. Only RMSEA values for both models are within acceptable range (i.e.,<.04).

Discussion

The aim of the present work was to translate Cornell Critical Thinking Test-Level Z into Urdu language, adapt it according to Pakistani culture and establish its psychometric properties. The literature search did not reveal any major indigenous measure of critical thinking abilities, so it was important to translate an already existing standardized measure in Urdu language for acceleration of critical thinking research in indigenous context. Cross-language validation study showed that participants got almost identical mean scores on both English and Urdu language versions of the test. Although most of the items showed non-significant item to item correlations, both versions showed a significant positive correlation between total scores. The participants have performed poorly on both tests in a cross-validation study and in a main study. Comparing with norms derived from various studies delineated in the test manual, the participants of the current study exhibited very low mean scores. The test manual reports mean difficulty indices ranging from 0.55 to 0.61 (a sample of undergraduate university students) (pp 15-19). The item analysis in the present study showed a rather high mean item difficulty and a lower mean discrimination index (point-biserial correlation) compared to the reported index in the manual. Taking many things into account— i.e. general educational environment of educational institutes in Pakistan, lack of exposure to critical thinking as part of curriculum, unfamiliarity of the students with such kind of testing, and the difficulty of the construct itself— the higher item difficulty and lower discrimination values are not surprising. In fact, it provides a glimpse into the state of critical thinking in our society.

To determine factor structure of the test, Confirmatory Factor Analysis (CFA) was performed using lavaan package (Yves, 2012) with R version 3.6.1 (R Core Team, 2019). Due to dichotomous nature of test items, a more suitable Diagonally Weighted Least Squares (DWLS) method of estimation was used to see model fit and calculate coefficients (Rhemtulla et al., 2012). Several recommended model fit

indices were used to observe model fit (Byrne, 2016). A theoretically predicted 5-factor model did not fit well with all indices except RMSEA. A 4-factor model was run after removing items from induction subtest as these items were showing negative loadings on its factor. Though, the 4-factor solution also does not seem appropriate.

Factorial validity is a great concern for Cornell tests and no appropriate factor structures for these tests have yet been obtained. Perhaps, due to complex nature of construct with overlapping of subtests, there is limited empirical evidence available on its dimensionality (Leach et al., 2020). There are some published studies that have tried to determine factorial sense of CCTT across different cultures. But most of these studies have achieved discrepant findings to reach at a simple theoretical structure (Brown, 2006; Follman et al., 1970; Frisby, 1992; Gross, 1996; Leach et al., 2020; Michael et al., 1980; Verburgh et al., 2013). In one of the earliest studies, Michael et al. (1980) used principal component analysis on level X of the test and discovered only Deduction subtest to be theoretically compatible. Verburgh et al. (2013), in a Dutch version of CCTT- level Z, found a unidimensional structure with five subscales correlating sufficiently to form a single dimension by using multiple item response theory (MIRT). Leach et al. (2020) applied EFA as well as CFA to establish structural dimension of CCTT-level X that also resulted in reduction of items and fewer interpretable factors.

The authors of the test are also cognizant of this structural complexity and have discussed this interdependence and overlap among multiple dimensions of critical thinking in these words.

"…Putting all this together, one can see the pervasiveness of basic deduction, the perhaps even greater pervasiveness of ability to deal with meaning, and the consequent theoretical difficulty of securing totally independent part scores and strong independent factors in a factor analysis. Critical thinking is not a unidimensional concept either, making it difficult to obtain high internal consistency reliability estimates" (Ennis et al., 2005, p. 3).

The available empirical evidence and test developers' own opinion lead to a conclusion that a suggested factor structure cannot be taken as evidence for test validity. Somehow, if that structure is obtained in some very sophisticated settings, this would not be a universal one and might show many cross-cultural inconsistencies as evidenced in various studies already discussed.

The test's inability to have meaningful factorial dimensions cannot mean that the test is completely invalid. Ennis et al. (2005) have presented various other evidence of test's criterion and content validity (pp. 20-21). CCTT tests are based on Cornell/Illinois model, a critical thinking framework that is developed after years of work in that particular domain (Ennis, 2011b, 2015b). The test developers have informed that, while developing this instrument, the CCTT items have been thoroughly discussed among members of the Illinois Critical Thinking Project and a universal agreement among the experts, on the correctness of the keyed answers was achieved. The test is claimed to have strong content validity as ample expert agreement exists about relevance of its items with Cornell/Illinois framework.

The test scores are somewhat more interpretable considering the point-biserial correlation coefficients. Twelve items showed non-significant correlation coefficients. The Cronbach alpha reliability coefficient, for the full-length test, improved from α=.47 to α=.59 after removing these items from the test. The test manual also mentions various studies on CCTZ with its reliability coefficients ranging between .49 to .87 (Ennis et al., 2005, p. 17). This alpha value for the current translated version seems reasonable enough for the test that seems neither unidimensional nor has an unambiguous multidimensional structure. As the test is measuring multiple aspects within it, so very high overall internal consistency value

for CCTT-Z is neither possible nor desirable.

Conclusion

The English and Urdu versions show positive correlation indicating similarity of construct. The difficulty level of the test is very high, and the test has poor factor structure. Several other studies have failed to identify a meaningful factor structure and the test mostly relies on content related and convergent validity evidence. It seems that the Urdu version of the test is useful once Induction subtest items are removed which leads to improved reliability scores. As both Urdu and English language versions seem compatible, and as English version provides ample evidence of test's construct validity, future studies need to use Urdu version along with other relevant measures of critical thinking, achievement, and intelligence to further expand on the convergent validity of the Urdu version.

Contribution of Authors

Abdul Wahab Liaquat: Conceptualization, Investigation, Data Curation, Formal Analysis, Writing - Original draft

Najam ul Hasan Abbasi: Conceptualization, Methodology, Writing-Reviewing & Editing, Supervision

Siara Ferdous: Conceptualization, Methodology, Investigation, Writing – Original Draft

Conflict of Interest

There is no conflict of interest declared by authors.

Source of Funding

The authors declared no source of funding.

References

Akhtar, F. (2019). Critical thinking in education: An answer to extremism. *ISSRA Papers, 11*(1), 21-38.

Anderson, M. (2015). The case against" critical thinking skills": In pursuit of a humanizing pedagogy. *Philosophical Studies in Education, 46*, 83-89.

Barnouw, V. (1982). *An Introduction to Anthropology: Ethnology*. Homewood, IL: Dorsey Press.

Berry, J. (1969). On cross-cultural comparability. *International Journal of Psychology, 4*(2), 119-128.

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology, 1*(3), 187-196.

Brislin, R. W. (1986). The wording and translation of research instruments. In W. L. Lonner & J. W. Berry (Eds.), *Field Methods in Cross-Cultural Research* (pp. 137-164). CA, Newbury Park: Sage.

Brown, T. A. (2006). *Confirmatory Factor Analysis for Applied Research*. New York: The Guilford Press.

Byrne, B. M. (2016). *Structural Equation Modeling with AMOS: Basic Concepts, Application and Programming*. New York: Routledge.

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

Ennis, R. H. (1962). A concept of critical thinking. *Harvard Educational Review, 32*(1), 81-111.

Ennis, R. H. (1989). Critical thinking and subject specificity: Clarification and needed research. *Educational Researcher, 18*(3), 4-10. doi:10.3102/0013189x018003004

Ennis, R. H. (2002). An outline of goals and outcomes for a critical thinking curriculum and specifications for assessment. http://thehob.net/math/proces/reasngProf/thinkingCur.htm

Ennis, R. H. (2011a). Critical thinking: Reflection and perspective—part I. *Inquiry: Critical Thinking Across the Desciplines, 26*(1), 4-18.

Ennis, R. H. (2011b). The nature of critical thinking: An outline of critical thinking dispositions and abilities. In. MIT, Cambridge: Presentation at Sixth International Conference on Thinking.

Ennis, R. H. (2015a). Critical thinking: A Streamlined conception. In: Davies M., Barnett R. (eds) *The Palgrave Handbook of Critical Thinking in Higher Education*. Palgrave Macmillan, New York. https://doi.org/10.1057/9781137378057_2

Ennis, R. H. (2015b). The nature of crtical thinking: Outlines of general critical thinking dispositions and abilities. In. MIT, Cambridge: Presentation at Sixth International Conference on Thinking.

Ennis, R. H., & Millman, J. (1971). *The Cornell Critical Thinking Test*. Urbana, Illinois: Critical Thinking Project.

Ennis, R. H., Millman, J., & Tomko, T. (1985). *Cornell Crtical Thinking Tests Level X and Level Z-Manual*. Pacific Grove, CA: Midwest Publishing.

Ennis, R. H., Millman, J., & Tomko, T. N. (2005). *Cornell Critical Thinking Tests: Administration Manual*: The Critical Thinking Co.

Facione, P. A. (1990). *Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction (The Delphi Report)*: The California Academic Press.

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

Facione, P. A., Facione, N. C., Blohm, S. W., & Giancarlo, C. A. F. (1990). *The California Critical Thinking Skills Test: CCTST*. Millbrae, CA: California Academic Press.

Follman, J., Brown, L., & Burg, E. (1970). Factor analysis of critical thinking, logical reasoning, and English subtests. *The Journal of Experimental Education, 38*(4), 11-16.

Frisby, C. L. (1992). Construct validity and psychometric properties of the Cornell Critical Thinking Test (Level Z): A contrasted groups analysis. *Psychological Reports, 71*(1), 291-303. doi:10.2466/pr0.1992.71.1.291

Gross, M. L. (1996). Moral reasoning and ideological affiliation: A cross-national study. *Political Psychology, 17*(2), 317-338. doi:10.2307/3791813

Halpern, D. F. (2010). *Halpern Critical Thinking Assessment Manual*. Schuhfried GmbH.

Halpern, D. F. (2013). *Thought and Knowledge: An Introduction to Critical Thinking*: Psychology Press.

Haseli, Z., & Rezaii, F. (2013). The effect of teaching critical thinking on educational achievement and test anxiety among junior high school students in Saveh. *European Online Journal of Natural and Social Sciences, 2*(2), 168-175.

Hirayama, R., Tanaka, Y., Kawasaki, M., & Kusumi, T. (2010). A study on the composition and nature of the Japanese version of the critical thinking scale : Using Cornell Critical Thinking Test Level Z. *Journal of the Japanese Society of Educational Engineering, 33*(4), 441-448.

Kaplan, R. M., & Saccuzzo, D. P. (2005). *Psychological Testing: Principles, Applications, and Issues*: Wadsworth.

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

McPeck, J. (1981). *Critical Thinking and Education*. London: Routledge.

Michael, J. J., Devaney, R. L., & Michael, W. B. (1980). The factorial validity of the Cornell Critical Thinking Test for a junior high school sample. *Educational and Psychological Measurement, 40*(2), 437-450. doi:10.1177/001316448004000223

Nauman, S. (2017). Lack of critical thinking skills leading to research crisis in developing countries: A case of Pakistan. *Learned Publishing*, 1-4. doi: 10.1002/leap.1091

Phillips, L. R., & Luna, I. (1996). Toward a cross-cultural perspective of family caregiving. *Western Journal of Nursing Research, 18*(3), 236-252.

Rashid, S. & Qaisar, S. (2017). Developing critical thinking through questioning strategy among fourth grade students. *Bulletin of Education and Research*, 38(2), 153-168.

R Core Team (2019). R: A language and envrironment for statistical computing. In. Vienna, Austria: R Foundation for Statistical Computing.

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. doi:https://doi.org/10.1016/j.intell.2020.101487

Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods. *Psychological Methods, 17*, 354-373.

Riaz, M. N. (2017). *Test Construction: Development and standardization of psychological tests in Pakistan*. Islamabad: IRD, International Islamic University.

Sahin, H., French, B. F., Hand, B., & Gunel, M. (2015). Detection of differential item functioning in the Cornell Critical Thinking Test between Turkish and United States students. *European Journal of Psychological Assessment, 31*(4), 238-246. doi:10.1027/1015-5759/a000230

Sanders, M., & Moulenbelt, J. (2011). Defining critical thinking: How far have we come? *Inquiry: Critical Thinking Across the Disciplines, 26*(1), 38-46.

Shahzadi, U., Nimmi, & Khan, I. (2020). Exploring the relationship between critical thinking skills and academic achievement. *Sir Syed Journal of Education and Social Research, 3*(1), 236-242. doi:https://doi.org/10.36902/sjesr-vol3-iss1-2020(236-242)

Shirazi, F., & Heidari, S. (2019). The relationship between critical thinking skills and learning styles and academic achievement of nursing students. *The Journal of Nursing Research, 27*(4), e38-e38. doi:10.1097/jnr.0000000000000307

Taghva, F., Rezaei, N., Ghaderi, J., & Taghva, R. (2014). Studying the relationship between critical thinking skills and students' educational achievement (Eghlid universities as case study). *International Letters of Social and Humanistic Sciences, 25*, 18-25. doi:10.18052/www.scipress.com/ILSHS.25.18

Verburgh, A., François, S., Elen, J., & Janssen, R. (2013). The assessment of critical thinking critically assessed in higher education: a validation study of the CCTT and the HCTA. *Education Research International, 3*. doi: 10.1155/2013/198920

Watson, G. B., & Glaser, E. (2007). *Watson - Glaser Critical Thinking*

Appraisal. Frankfurt am Main: Harcourt Test Services.

Yves, R. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1-36.