Kurdish Studies
September 2024
Volume: 12, No: 5, pp 1041-1048
ISSN: 2051-4883 (Print) | ISSN 2051-4891 (Online)
www.KurdishStudies.net

DOI: 10.53555/ks.v12i5.3400

Unlocking Intelligence Across Languages: The Urdu Validation Of The Slosson Intelligence Test (SIT-4) Comprehension Domain For University Students

Uzma Qayyum Khan[1], Muhammad Faran, Ph.D[2*], & Noshi Iram Zaman, Ph.D[3]

[1]Department of Professional Psychology, Bahria University, Islamabad, Pakistan. Email: uzmaqkhan@gmail.com, Orcid: 0000-0003-0750-9910
[2*] Department of Professional Psychology, Bahria University, Islamabad, Pakistan. Email: mfaran.buic@bahria.edu.pk, Orcid: 0000-0003-1200-8846
[3] Department of Professional Psychology, Bahria University, Islamabad, Pakistan. Email: noshi.zaman@bahria.edu.pk, Orcid: 0000-0002-6556-8030

***Corresponding Author:-** Muhammad Faran, Ph.D
[*] Department of Professional Psychology, Bahria University, Islamabad, Pakistan. Email: mfaran.buic@bahria.edu.pk, Orcid: 0000-0003-1200-8846

Abstract
This research aimed to translate and validate the Comprehension Domain of the Slosson Intelligence Test-4 (SIT-4) into Urdu for university students in Pakistan. The translation followed a standardized forward and backward process, followed by a pilot study. A confirmatory factor analysis was performed on a sample of 900 students (427 males, 473 females) with an average age of 24.38 years (SD = 3.45). The results indicated strong psychometric properties, demonstrating validity and reliability for both male and female participants. The Urdu version of the SIT-4 (comprehension domain) was found to be culturally and linguistically appropriate for this population.

Introduction
As evidenced across various disciplines, culture exerts a substantial influence on an person's critical thinking abilities and intelligence level, rendering intelligence a culturally contingent concept (Panicker et al., 2006). The effectiveness of a person's performance on a particular task is shaped by the specific requirements of the task, as well as the individual's background knowledge and characteristics. Intelligence testing serves several educational and therapeutic purposes, emphasizing the necessity of a comprehensive and well-standardized assessment method. Evaluating intellectual capacity is crucial for predicting learning potential and holds substantial value in both clinical and research settings (Ambreen & Kamal, 2014).
Given the multicultural and multilingual nature of contemporary societies, assessments are routinely adapted to facilitate cross-cultural comparisons. It is imperative to adhere to best practices during test modification to ensure accuracy and comparability of results. The translation and adaptation process should accurately maintain the content and cultural significance of the original and modified tests, ensuring that results can be compared across diverse cultures (Hernández et al., 2020).
The linguistic validation of the Comprehension Domain of the Slosson Intelligence Test- 4th edition (SIT-4) in Pakistan's native language (Urdu) was considered essential. Intelligence testing has garnered increased interest in Pakistan for its potential in research, education, and scientific advancement. Despite its significance, there is a scarcity of studies focusing on test formulation, with previous attempts often marred by methodological and content flaws, making them inadequate (Hussain, 2001). SIT-4, designed as a brief screening test to assess general verbal cognitive ability, faces limitations in countries where English is not the primary language due to linguistic barriers and content imbued with cultural meaning. The Comprehension Domain has been defined by Richard Slosson as the cognitive ability that reflects an individual's grasp of social norms, practical wisdom, and ability to comprehend proverbs and idioms. It showcases awareness of social cues, adeptness in appropriate behavior in various situations, and strategies for navigating social challenges. In essence, it exemplifies aspects of social intelligence (Erford et al., 2017). Slosson Intelligence Test (SIT-4) comprises six cognitive domains, heavily laden with verbal elements that signal culturally and linguistically unsuitable content (Aziz & Ahmad, 1993). SIT-4 also assesses crystallized intelligence, which is greatly influenced by cultural factors (Weiss et al., 2006).
Rather than developing a new test from scratch, researchers often opt to translate and modify existing validated tests from other cultures, saving time and resources while facilitating cross-cultural comparisons (Hambleton & Lee, 2013). Given the prevalence of diverse and multilingual settings, this approach is considered crucial (Duarte & Rossier, 2008). Thus, the impetus for the translation and linguistic validation of this test arose from the dire need for an intelligence screening tool in Pakistan's native language. The translation was conducted in Phase I of the study, and Phase II of the study focused on the linguistic validation.

Method

Research Design

The research was conducted in two distinct phases, following a cross-sectional correlation design. During the first phase (Phase I), the comprehension domain of the SIT-4 was translated using the standardized forward-backward translation method (WHO, 2020). In the second phase (Phase II), the linguistic validation of the translated comprehension domain was carried out. Authorization for both the translation and validation processes was secured from the original test author.

Phase I – Phase I Translation of Comprehension Domain of SIT- 4.

Beaton's et al., (2000) approach was followed for cross-cultural adaptation and translation of SIT-4.

Step 1 - Forward Translation: Five bilingual professionals were tasked with translating the Comprehension Domain of SIT-4. This group consisted of three PhD candidates in Psychology and two individuals with Master's degrees in the field. They were chosen specifically for their knowledge of test development and their awareness of cultural nuances. The translators were directed to maintain the original meaning and difficulty of the items while ensuring that the translated text was both clear and easy to understand.

Step 2- Backward Translation: Once the Urdu forward translation was finalized, a separate group of three bilingual experts, all holding PhDs in Psychology, conducted the reverse translation back into English. This backward translation aimed to identify any inconsistencies or contextual discrepancies by comparing the original English version with the back-translated material. Importantly, these experts were independent of the forward translation process and had no prior exposure to the original test. After completing the three separate back-translations, the researcher and supervisor carefully examined and compared them to identify any differences or commonalities. Ultimately, a single backward translation was selected, and when compared to the original test, no significant differences were found.

Step 3 – Review and Scrutiny: After finalizing and making necessary adjustments to the Urdu version of the Comprehension Domain of SIT-4, the final phase involved a detailed review for proofreading and grammatical accuracy. This rigorous examination revealed no significant errors or inconsistencies.

Step 4 – Cognitive Debriefing: The objective in this phase was to assess the understanding and clarity of Urdu translation for comprehension domain of SIT-4. This evaluation was performed in a group of fifty participants (19 to 28 years; both genders). The participants were briefed about the purpose and procedure prior to the administration. The procedure for administering the test took place uniformly. Focus groups were discussed and interviews sought opinions on cultural equivalence, item difficulty and instructions clarity. The scoring was carried out as per SIT-4 Technical and Interpretative Manual (2017) guidelines. The subtest reliability was estimated by the Kuder-Richardson (KR-20) method, yielding a Reliability Coefficient Rho of. 79.

Step 5 – Final Version: After the cognitive debriefing, changes were made in order to improve the intelligibility, cultural appropriateness and correctness of the SIT-4 Comprehension Domain's Urdu translation. The revised version was reviewed by an expert committee to evaluate whether the changes made still convey the overall meaning and purpose of the test items. After this review was finished, the test was deemed fit for use with the larger group of individuals it had been intended for. The last version was tested in order to check the linguistic and culture appropriateness of the test and preserve its psychometric properties.

Phase II: Validation for Comprehension Domain of SIT-4

The phase II aimed to validate the Urdu version of the SIT-4 comprehension domain, focusing on its reliability and validity (Lamers et al., 2011). Both male and female participants were included. To ensure a comprehensive validation, a purposive sample of 900 individuals, aged 19 to 28, was selected from different educational institutions in Islamabad and Rawalpindi, Pakistan. At this point, the psychometric characteristics of the Urdu-translated comprehension domain were rigorously evaluated, examining its factor structure, reliability, and validity, with special attention to convergent validity.

Research Sample

For the empirical evaluation of the SIT-4 comprehension domain, 900 students were selected, consisting of 427 men and 473 women, with the mean age of 22.38 and standard deviation was 3.11. The inclusion criteria for the study required participants to be between 19 and 28 years old and currently enrolled in an educational institution. Individuals with visual impairments, learning disabilities, exceptionalities, or intellectual disabilities were excluded (Slosson, 2017). The sample size followed the guidelines by Tabachnick and Fidell (2013), recommending a 10:1 ratio, meaning at least 10 participants per item. A convenience sampling technique was used to select participants from various institutions in Islamabad and Rawalpindi, Pakistan.

Research Instrument and Procedure

The intelligence test, originally developed by Richard L. Slosson (Slosson, 1963), is comprised of six distinctive cognitive domains. These domains included, vocabulary, general information, comprehension, similarities and differences, quantitative, and auditory memory.

The comprehension domain evaluates an individual's understanding of societal norms, practical wisdom, and ability to interpret proverbs and expressions. It highlights awareness of social cues, the ability to behave appropriately in different contexts, and strategies for addressing social challenges, reflecting elements of social intelligence. However, knowing the right social behaviors doesn't always translate into consistently following social norms. The comprehension tasks assess a range of cognitive skills such as reading comprehension, reasoning, critical thinking, problem-solving, and inference-making, offering

insights into how well individuals understand and apply information. This domain has long been a key feature of most comprehensive intelligence tests.

The SIT-4 consists of 187 items, with 33 items dedicated to the Comprehension Domain. Scoring for the test is straightforward, with correct answers receiving a score of 1 and incorrect answers receiving 0. The manual provides clear and easy-to-follow instructions for evaluation. Scoring is based on identifying a basal score, achieved through ten consecutive correct responses, and a ceiling score, marked by ten consecutive incorrect responses. Testing stops once the ceiling score is reached. The total raw score is determined by counting all correct answers from the basal to the ceiling level, with additional credit given for items answered correctly beyond the basal score (Slosson, 2017).

Data Analysis

The data was analyzed into two key steps, in step one, descriptive statistics including frequencies percentages, mean and mean, standard deviation were calculated for participant's characteristics. Additionally, confirmatory factor analysis (CFA) was employed through structural equation modeling (SEM) to rigorously assessing the psychometric properties of the Urdu-adapted Comprehension Domain of the SIT-4, with a particular focus on its application to university student populations. However, for the conceptual equivalence of the measure across men and women, measurement invariance was also conducted.

Results

Table 1 presents the descriptive statistical analysis of the characteristics of the participants. The study sample consisted of 47.33% males and 52.66% females, with an average age of 22.83 years and a standard deviation of 3.1 years. Around 46.11% of the participants were attending public sector institutions, while 53.88% were enrolled in private sector institutions. Furthermore, 71.88% of the respondents were undergraduate students, and 28.11% were pursuing postgraduate degrees.

Table 1 *Descriptive Statistics of the Participants Characteristics (N = 900)*

Variables	***f(%)***	***M(SD)***
Age (years)		22.83 (3.11)
Gender		
Men	427 (47.33)	
Women	473 (52.66)	
Institute		
Public	415 (46.11)	
Private	485 (53.88)	
Education		
Undergraduate	647 (71.88)	
Postgraduate	253 (28.11)	

A Confirmatory Factor Analysis (CFA) was conducted using Structural Equation Modeling (SEM) to validate the factor structure of the Urdu-translated version of the SIT-4 comprehension domain. This analysis was applied to the 33 comprehension items within the domain. The indices of the model fit for the proposed test structure are detailed in Table 2.

Table 2 *Confirmatory Factor Analysis, Fit Indices for Comprehension Domain of SIT-4 for Adults (N = 900).*

Model	χ^2	*df*	χ^2/df	*GFI*	*CFI*	*NFI*	*RMSEA*	*SRMR*
Initial Model	3694.38	990	3.73	.84	.83	.81	.07	.06
Model Fit	1984.74	986	*2.01*	.94	.93	.92	.06	.05
$\Delta\ \chi^2$	*1709.64**							

Note. GFI= Goodness of fit index, CFI=comparative fit index, NNFI = non-normed fit index; RMSEA=root mean square error of approximation, SRMR=Standardized root means square, $\Delta\chi^2$ = chi-square change.

Table 2 also highlighted the fit indices for the comprehension domain to measure the absolute and relative fitness of the fitted model. The absolute fit indices of initial model indicated poor fit with χ^2 (986) = 1984.74, $p < .05$. The chi-square statistic, which measures absolute model fit, is often influenced by sample size and the number of parameters estimated in a model (Hair et al., 2010). Therefore, it is recommended to also consider relative fit indices, such as the GFI, CFI, NFI with the null model comparison, and RMSEA, and SRMR for saturated model comparison.

Hu and Bentler (1999), along with other theorists, posited that an optimal model should exhibit a χ^2/df ratio within the range of 0 to 3. Moreover, RMSEA and SRMR values should not exceed .08, while CFI, NFI, and GFI values are expected to be .90 or above. However, the initial fit indices for the model did not align with these benchmarks, necessitating modifications. Covariances were introduced between error terms, selected for their contextual relevance as indicated by the comprehension subtest's measurement model (Kenny, 2011). Following the integration of these covariances, the relative fit indices were recalculated. The adjusted GFI, CFI, and NNFI values reached .94, .93, and .92, respectively, while RMSEA and SRMR values were reduced to .06 and .05, reflecting an excellent model fit.

Confirmatory Factor Analysis (CFA) was employed to investigate the underlying factor structure of the Comprehension Domain of SIT-4. Once the model fit criteria were satisfactorily met, the analysis proceeded to assess reliability (internal consistency) and convergent validity within the sample. As per recommendations from Hair et al. (2010) and Henseler et al.

(2016), composite reliability and Cronbach's alpha should ideally surpass a threshold of .70, whereas the index of convergent validity, i.e., average variance extracted (AVE) should be .50 or higher to confirm convergent validity. The Comprehension Domain of SIT-4 demonstrated robust convergent validity, with AVE values of 0.51 for male participants and 0.52 for female participants. Furthermore, composite reliability and Cronbach's alpha coefficients ranged between 0.91 and 0.92, reflecting excellent reliability levels.

Figure 1 *Confirmatory Factor Analysis of Comprehension (Domain) of SIT -4 Urdu for Adults (N = 900).*

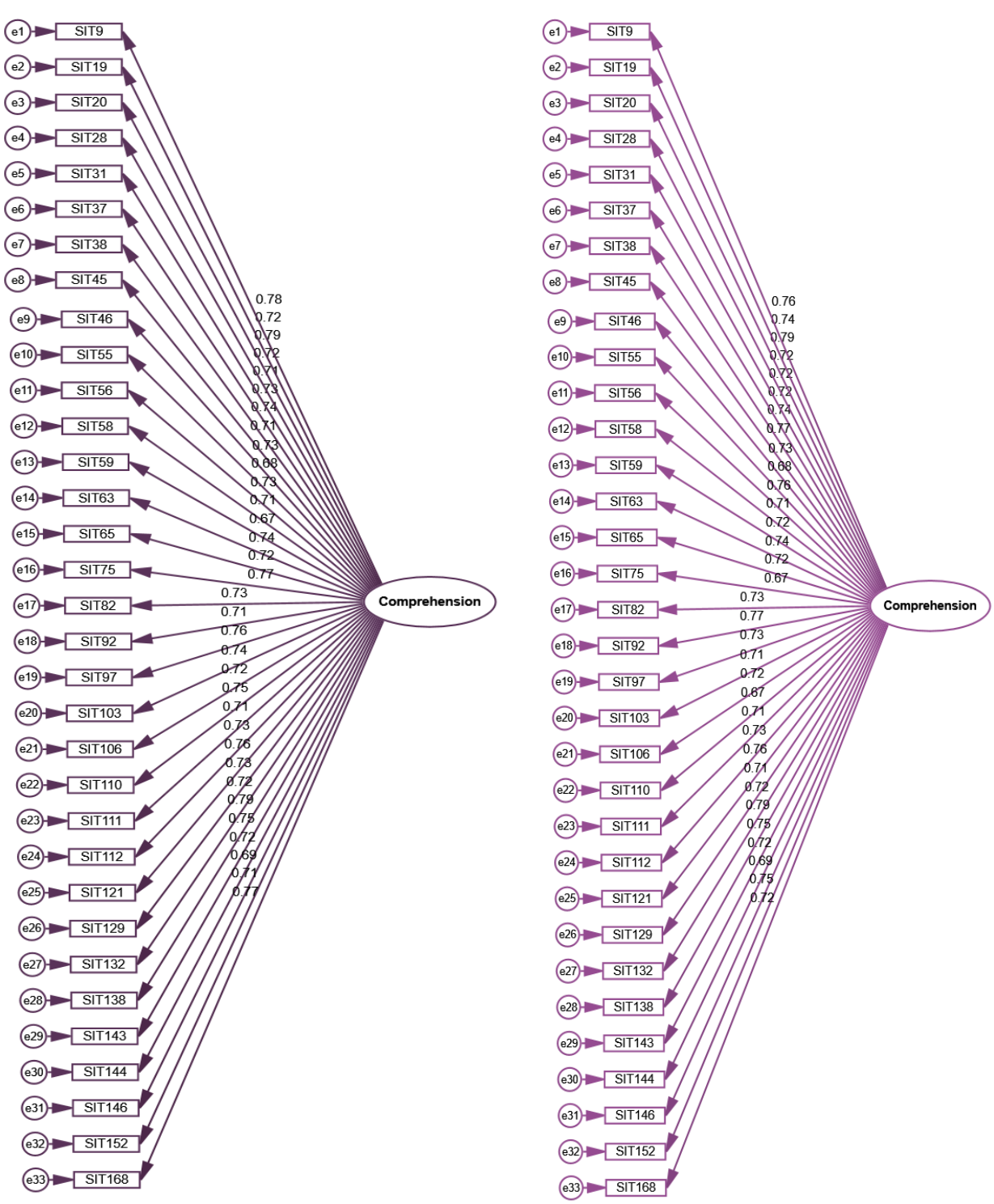

Table 3 *Psychometric Evaluation of Comprehension (Domain) of SIT -4 Urdu for University Students (N = 900).*

Items	Men				Women			
	α	*CR*	*AVE*	λ	*α*	*CR*	*AVE*	λ
Comprehension	.92	.94	.54		.90	.92	.53	
SIT 9				0.78				0.76
SIT 19				0.72				0.74
SIT 20				0.79				0.79
SIT 28				0.72				0.72
SIT 31				0.71				0.78
SIT 37				0.73				0.72
SIT 38				0.74				0.74
SIT 45				0.71				0.77
SIT 46				0.73				0.73
SIT 55				0.68				0.68
SIT 56				0.73				0.76
SIT 58				0.71				0.71
SIT 59				0.67				0.72
SIT 63				0.74				0.74
SIT 65				0.72				0.72
SIT 75				0.77				0.67
SIT 82				0.73				0.73
SIT 92				0.71				0.77
SIT 97				0.76				0.73
SIT 103				0.74				0.71
SIT 106				0.72				0.72
SIT 110				0.75				0.67
SIT 111				0.71				0.71
SIT 112				0.73				0.73
SIT 121				0.76				0.76
SIT 129				0.73				0.71
SIT 132				0.72				0.72
SIT 138				0.79				0.79
SIT 143				0.75				0.75
SIT 144				0.72				0.72
SIT 146				0.69				0.69
SIT 152				0.71				0.75
SIT 168				0.77				0.72

Note. CR = Composite reliability, AVE = Average variance extracted, λ (Lambda)= standardized factor loading

Measurement Invariance of Comprehension (Domain)

Measurement invariance was carried out across both male and female participants to evaluate whether individuals from distinct demographic groups perceive and interpret the assessment in a comparable manner, both conceptually and contextually (Byrne & Van de Vijver, 2010). There are two main types of measurement invariance due to the apparent existence of measurement invariance: full measurement invariance and partial measurement invariance which comprise of various forms such as configural, metric, scalar factor variance-covariance, and error variance invariance. This study focused on testing configural, metric, and scalar invariance to evaluate partial invariance (Hair et al., 2010). Measurement invariance was assessed using a nested modeling approach, starting with a configural model that did not impose constraints and provided a good fit for the entire dataset without dividing it by groups (refer to Table 4).

Table 4 *Testing for Measurement Invariance of Comprehension (Domain) of SIT -4 Urdu for University Students (N = 900).*

Model	χ^2	*df*	$\Delta\chi^2$	Δ*df*	*CFI*	Δ*CFI*	*RMSEA*
Unconstrained	1452.93	493	-		.94	-	.07
Configural Invariance	1984.74	986	531.81	493	.945	.005	.067
Matric Invariance	2028.45	1019	43.71	33	.948	.003	.063
Scalar Invariance	2070.67	1052	42.22	33	.952	.005	.056
Factor Covariance Invariance	-	-	-	-	-	-	-
Error Variance Invariance	2114.22	1085	43.55	33	.958	.006	.051

Note. CFI=comparative fit index,; RMSEA=root mean square error of approximation, SRMR=Standardized root mean square, $\Delta\chi^2$ = chi square change, Δdf= degree of freedom change, Δ*CFI*= comparative fit index change.

A comparative analysis was conducted to evaluate parameter invariance across restricted models for different groups, i.e., men and women. This method of assessing measurement invariance involves sequentially comparing constrained models to the unconstrained model. As invariance levels are identified, constraints are gradually introduced, and the models are evaluated in stages. Theorists suggest using the likelihood ratio test or chi-square difference test to assess the fit of nested models (Hair et

al., 2010). If the chi-square difference test is not statistically significant ($p > .05$), it indicates that the models are invariant across groups.

However, the chi-square test is known to be sensitive to factors such as sample size, the number of factors, and non-normal distribution (Hair et al., 1999). Cheung and Rensvold (2002) recommend using the Cumulative Fit Index (CFI) as a more reliable criterion for assessing invariance, where a change of 0.01 or less in the CFI is acceptable when applying constraints to nested models.

Initially, an unconstrained (unsplit) model was compared to a well-fitting multi-group (constrained) model. The results indicated $\Delta\chi^2 = 531.81$ with $\Delta df = 493$ at $p > .05$, and ΔCFI was .005, suggesting that the factor structure of the Comprehension Domain of SIT-4 was invariant between men and women. Next, metric invariance, which tests equal factor loadings and is crucial for measurement invariance, was analyzed. As detailed in Table 3, when factor loadings for the Comprehension Domain of SIT-4 were constrained between genders, $\Delta\chi^2 = 43.71$ with $\Delta df = 33$, and the CFI difference remained within the 0.01 range, confirming metric invariance between males and females.

The third phase involved evaluating scalar invariance, which examines the equality of means and intercepts across groups. Scalar invariance was confirmed, as both the metric and scalar models had CFI changes below the 0.01 threshold, with $\Delta\chi^2 = 42.22$ and $\Delta df = 33$, indicating that the means and intercepts were invariant across genders. Additionally, evidence of error variance invariance showed that both men and women exhibited similar error variances, with $\Delta CFI = .006$ and $\Delta\chi^2 = 43.55$ with $\Delta df = 33$ at $p > .05$. These findings confirm that the comprehension domain of SIT-4 is partially invariant across both genders.

Discussion

In recent years, the prevalence of test adaptations and validation has significantly increased and is now recognized as recommended practice (Epstein, et al, 2015; Hambleton & Zenisky, 2011). Transcultural adaptation refers to modifying an assessment tool to ensure it maintains its original validity while accounting for linguistic and cultural differences (Valer et al., 2015). In the initial phase of this study, three items from the comprehension domain of the SIT-4 were translated into clear and accessible Urdu, with the primary goal of ensuring contextual relevance and clarity.

In the second phase, linguistic validation was conducted with both male and female participants. Evaluating the psychometric properties of a test is vital in any applied field, as its validity and reliability determine its strength. The psychometric assessment of the comprehension domain demonstrated high reliability, with composite reliability and Cronbach's alpha coefficients ranging from .91 to .92. Additionally, the construct's convergent validity, measured by the average variance extracted (AVE), yielded excellent results according to the suggested threshold of measurement validations (Hair et al., 2010; Henseler et al., 2016).

To confirm the model's fit, a modification process was applied, resulting in fit indices that were deemed excellent. Furthermore, a measurement invariance test was conducted to evaluate the scale's consistency across genders. Measurement invariance is a necessary precondition for being able to accurately compare results across groups of interest (for example, gender or age groupings) because it assures that the concept included in measurement really contains its interpretation holds equally well within different subgroups. Measurement invariance is a prerequisite for identifying if the observed group differences on constructs reflect real or just procedural /score differences due to how groups understand and respond to measures (Cheung & Rensvold, 2002).

The findings of measurement invariance indicating that configural and metric invariances were achieved for both men and women. The configural invariance implies that for a measurement model the overall factor structure is identical across groups. Specifically, this is saying that the way observed variables are related to their latent construct (e.g., satisfaction or intelligence) should be consistent across groups. But if you don't have configural invariance, then one cannot be certain that the same construct is being measured similarly across any comparison groups (Meredith, 1993). However, Byrne and Van de Vijver (2010) suggested that metric invariance ensures that comparisons between groups are accurate by confirming that the relationship between the items and the underlying concept is the same across groups.

Lastly, the assessment of mean and intercept invariance (scalar invariance) confirmed that the comprehension domain of the SIT-4 is partially invariant across genders. Putnick and Bornstein (2016) argued that scalar invariance tests whether item intercepts are the same across groups, ensuring that differences in scores reflect true differences in the construct, not variations in how groups interpret the scale. Without it, comparing group means could lead to biased results.

Conclusion

The SIT-4 is an intelligence screening tool designed for both children and adults, requiring no specialized training and typically taking 15-20 minutes to administer and score. This study focused on translating and adapting the Comprehension section of the SIT-4 into Urdu, making it accessible for use in non-English speaking regions such as Pakistan, where Urdu is the national language. The findings indicated that the Urdu-adapted items in the comprehension domain demonstrated satisfactory validity and reliability, with strong psychometric properties for both men and women. A measurement invariance test confirmed the conceptual equivalence of the SIT-4 across genders. Moreover, the results suggest that the Urdu version of the SIT-4 comprehension domain holds promise for future research in Pakistan. Overall, this adaptation could enhance intelligence screening practices among university students in the country.

Limitations and Recommendations

This study encountered several limitations. Future research utilizing the adapted Comprehension Domain of the SIT-4 should aim to include participants from a broader age range and more diverse geographic regions to ensure a more comprehensive and representative sample. Additionally, incorporating a wider array of demographic variables, such as age, educational background, socioeconomic status, and parental education, could offer deeper insights into the psychometric characteristics of the adapted SIT-4. Examining these factors could reveal potential influences on test performance and validity, further enhancing the tool's applicability across different subgroups. Future studies could also explore longitudinal designs to assess the stability and predictive validity of the adapted Comprehension Domain over time.

References

1. Ambreen, S., & Kamal, A. (2014). Development of Norms of the Adapted Verbal Comprehension Index Subtests of WISC-IV for Pakistani Children. *Journal of Behavioural Sciences, 24*(1), 85-97. http://pu.edu.pk/images/journal/doap/PDFFILES/Abstract%20No%207_V_24_No_1_%2 02014.pdf
2. Aziz, S., & Ahmad, I. (1993). Adaptation of columbia mental maturity scale in pakistan. *Pakistan Journal of Psychological Research, 8*(1-2), 31-42. https://pjpr.scione.com/cms/abstract.php?id=568
3. Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine, 25*(24), 3186–3191. https://doi.org/10.1097/00007632-200012150-00014
4. Byrne, B. M., & Van de Vijver, F. J. (2010). Testing for measurement and structural equivalence in large-scale cross-cultural studies: Addressing the issue of nonequivalence. *International Journal of Testing, 10*(2), 107-132. https://doi.org/10.1080/15305051003637306
5. Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural equation modeling, 9*(2), 233-255. https://doi.org/10.1207/S15328007SEM0902_5
6. Duarte, M. E., & Rossier, J. (2008). Testing and assessment in an international context: Cross-and multi-cultural issues. In *International handbook of career guidance* (pp. 489-510). Springer. https://doi.org/10.1007/978-1-4020-6230-8_24
7. Erford, B. T., Larson, S. L., & Slosson, S. W. (2017). *Slosson Intelligence Test – 4th edition for children and adults: Administration manual.* Slosson Educational Publications.
8. Epstein, J., Santo, R. M., & Guillemin, F. (2015). Cross-cultural adaptation of questionnaires: review of concepts and current guidelines. *Journal of Clinical Epidemiology, 68*(4), 435- 441. https://doi.org/pi.2014.11.021
9. Hair, J. D., Anderson, R. E., Tatham, R. L., & Black, W. C. (2010). *Multivariate date analysis* (7th ed.). New Jersey: Prentice Hall.
10. Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, C. B. (1999). *Análisis Multivariante*. Pearson Prentice Hall. https://dialnet.unirioja.es/servlet/libro?codigo=320227
11. Hambleton, R. K., & Lee, M. K. (2013). Methods for translating and adapting tests to increase cross-language validity. *The Oxford handbook of child psychological assessment*, 172-181. https://doi.org/10.1093/oxfordhb/9780199796304.013.0008
12. Hambleton, R. K., & Zenisky, A. L. (2011). Translating and Adapting Tests for Cross- Cultural Assessments. In D. Matsumoto & F. J. R. van de Vijver (Eds.), *Cross- Cultural Research Methods in Psychology* (pp. 46-74). Cambridge University Press. https://doi.org/10.1017/CBO9780511779381.004
13. Henseler, J., Hubona, G., & Ray, P. A. (2016). Using PLS path modeling in new technology research: updated guidelines. *Industrial management & data systems, 116*(1), 2-20. https://doi.org/10.1108/IMDS-09-2015-0382/full/html
14. Hernández, A., Hidalgo, M. D., Hambleton, R. K., & Gómez Benito, J. (2020). International test commission guidelines for test adaptation: A criterion checklist. *Psicothema, 32*(3), 390-398. https://doi.org/10.7334/psicothema2019.306
15. 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. https://doi.org/10.1080/10705519909540118
16. Hussain, S. S. (2001). *Development, validation, and standardization of a group verbal intelligence test in Urdu for adolescents.* (Unpublished Doctoral Dissertation). National Institute of Psychology, Quaid-i- Azam University, Islamabad, Pakistan.
17. Kenny, D. A. (2011). Respecification of latent variable models. *davidakenny.net, 12*(2012), 109- 119. http://davidakenny.net/cm/respec.htm
18. Lamers, S. M., Westerhof, G. J., Bohlmeijer, E. T., ten Klooster, P. M., & Keyes, C. L. (2011). Evaluating the psychometric properties of the mental health continuum-short form (MHC- SF). *Journal of clinical psychology, 67*(1), 99-110. https://doi.org/10.1002/jclp.20741
19. Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*(4), 525–543. https://doi.org/10.1007/BF02294825
20. Panicker, A. S., Hirisave, U., & Subbakrishna, D. (2006). WISC-III uk: Comparison of Indian and UK norms. *Journal of Indian Association for Child and Adolescent Mental Health, 2*(4), 108-111. https://doi.org/10.1177/0973134220060403
21. Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, *41*, 71–90. https://doi.org/10.1016/j.dr.2016.06.004
22. Slosson, R. (1963). *Manual: Slosson Intelligence Test for Children and Adults*. East Aurora, NY: Slosson Educational Publ. https://worldcat.org/en/title/25446906
23. Slosson, R. L. (2017). *(SIT-4) Slosson Intelligence Test-4th Edition.* Slosson Educational Publication. http://www.slosson.com/onlinecatalogstore_c394797.html
24. Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics* (6th ed.). Boston, MA: Pearson.

25. Valer, D. B., Aires, M., Fengler, F. L., & Paskulin, L. M. G. (2015). Adaptation and validation of the Caregiver Burden Inventory for use with caregivers of elderly individuals. *Revista Latino-Americana de Enfermagem, 23*, 130-138. https://doi.org/10.1590/0104-1169.3357.2534
26. Weiss, L. G., Saklofske, D. H., Schwartz, D. M., Prifitera, A., & Courville, T. (2006). Advanced clinical interpretation of WISC-IV Index Scores. In L. G. Weiss, A. Prifitera, D. H. Saklofske, & J. A. Holdnack (Eds.), *WISC-IV advanced clinical interpretation* (pp. 139-179). MA, California, London: Elsevier Inc. https://doi.org/10.1016/B978-012088763-7/50005-X
27. WHO. (2020). *Process of translation and adaptation of instruments*. World Health Organization. http://www.who.int/substance_abuse/research_tools/translation/en/