TPM Vol. 32, No. S3, 2025
ISSN: 1972-6325
https://www.tpmap.org/

Open Access

PSYCHOMETRIC VALIDATION OF THE URDU VERSION OF THE MODIFIED CHECKLIST FOR AUTISM IN TODDLERS (M-CHAT)

DR. KHIZRA MUSSADIQ
MBBS, FCPS PSYCHIATRY, NISHTAR MEDICAL UNIVERSITY & HOSPITAL, MULTAN, PAKISTAN,
EMAIL: khizramussadiq@gmail.com

DR. HAFIZ SHAFIQUE AHMAD
ASSOCIATE PROFESSOR PSYCHIATRY, D. G. KHAN MEDICAL COLLEGE, DERA GHAZI KHAN, PAKISTAN.
EMAIL: drhshafiq@ymail.com

DR. SADIA HADAYAT
MBBS, FCPS PSYCHIATRY, MUKHTAR A. SHAIKH HOSPITAL, MULTAN, PAKISTAN.
EMAIL: shadayat@gmail.com

DR. IRFAN SHABBIR
MBBS, FCPS PEDIATRICS, NISHTAR MEDICAL UNIVERSITY & HOSPITAL, MULTAN, PAKISTAN.
EMAIL: irfanshabbir151@gmail.com

ABSTRACT

The current study tried to carry out psychometrics validation of Urdu version of the Modified Checklist for Autism in Toddlers (M- CHAT) to construct a culturally balanced screening instrument in the early detection of autism in the Pakistani general populations. The study thus dealt with the significant absence of validated screening tools to detect autism in the Urdu-speaking community since all the existing screening tools could not be culturally and linguistically adapted in Pakistani milieu. The validation design used in the study is a cross-sectional design of 450 toddlers aged between 16-30 months, recruited in the sample at the pediatric clinics, child development centers, and community health facilities in Multan, Pakistan. The homegrown methodology entailed both forward and backward translation of the M-CHAT in the originating language, i.e., English to the target language i.e., Urdu, opinion of a panel of experts and pilot testing. The Urdu version of M-CHAT was administered to caregivers through the researcher's collection of data, including clinical evaluation of caregivers with the aid of standard diagnostic criteria as a gold standard. The statistical analysis involved reliability using Cronbach's alpha, construct validity via exploratory factor analysis and criterion validity by using sensitivity and specificity. The results showed an internal consistency (alpha 0.84) and reasonable test-retest reliability (r = 0.78). Urdu M-CHAT demonstrated good sensitivity (87%) and specificity (83%) in the detection of autism, the positive predictive value of 76 percent, and the negative predictive value of 91 percent. Factor analysis showed that the data could be partitioned into 3 factors that conformed to domains of core autism symptoms. The research findings stated that the Urdu M-CHAT was a valid and dependable screening scale in screening autism in Pakistani toddlers. The suggestions were the use of the approved tool in everyday pediatric practice, the education of health care professionals on early screening of autism, and the longitudinal study to determine the long-term duration of diagnosis and development outcome.

Keywords: Psychometrics, validation, Urdu version, Autism, Toddlers, screening, culturally, linguistically.

INTRODUCTION

Autism Spectrum Disorder (ASD) is perhaps one of the most complicated neurodevelopmental disorders that is surely going to affect children all over the world (Wang, Wang, Wu, Wang, & Sun, 2023). Autism is a disorder that persistently displays a deficiency of social communication and interaction coupled with restricted and repetitive patterns of behavior, interests, or activities. The rate of occurrence of autism has continued to rise all over the world with estimates indicating that about 1 out of every 36 children have ASD. This trend has been threatening and this is why healthcare systems are currently focusing on early detection and intervention solutions because every evidence has proved that early identification and intervention has a great advantage on the long-term prospects of the children with autism (Antolini & Colizzi, 2023). It is impossible to emphasize the importance of the early years of growth as the neural development and plasticity are at peak during their first years of life, hence, early spotting of any concern is critical for children with autism. Intensive behavioral interventions during this window can result in significant improvement in communication skills, social functioning and adaptive behaviors. Nevertheless, the issue is that the symptoms of autism can be too subtle in very young children, which is why they are quite difficult to detect early by the healthcare providers and other caregivers. The challenge is even complicated by developing nations whose awareness on autism is minimal; and cultural aspects can affect the view and reporting of the anxieties of development (Bonti, Zerva, Koundourou, & Sofologi, 2024).

In Pakistan, the case of autism detection and intervention is rather complicated. Although the healthcare system in the country is becoming better, it does not offer complete developmental screening programs, especially in the rural and semi-urban regions. Most children with autism are diagnosed well past the opportune time to intervene, and in many instances, it is not until the child gets to their school age that they acknowledge that this child has a problem that needs to be addressed. This slow diagnosis not only restricts the success of the interventions but also puts considerable emotional and financial pressure on the families who are no longer able to comprehend and assist their children in meeting their developmental needs (Salman, Arshad, & Sitwat, 2024). Many researchers have noticed that the screenings for understanding early autism have been done mostly by adopting the Modified Checklist for Autism in Toddlers (M-CHAT), due to its being one of the most validated autism screening tests in the world today. Initially, the M-CHAT questionnaire was modeled as a parent-reporting questionnaire to detect the risks of autism spectrum disorder in children aged 16-30 months. The instrument comprises 23 yes/no questions and evaluates many developmental milestones and patterns that apply to autism, like signs of social communication, joint attention, and repetitive functioning. It is popular because it is easy to administer, economical and it has shown reliability across various population groups (Bashir & Khanum, 2024).

In non-English speaking countries, there is a major difficulty in the application of such sound screening tools as the M-CHAT. The validity and reliability of screening tools may be influenced by matters of language, cultural outlooks on child-rearing procedures, and differences in the expectations of developmental milestones, none of which may be effectively transduced in the translation of the tool to be administered. It has been proved time and again that psychological and developmental assessment instruments need stringent validation procedures when translated to be used in various other cultures and languages to make them competent and suitable (Huda et al., 2024).

Urdu is a language that is spoken by more than 230 million people in the world, and it is the national language of Pakistan, which is a very large linguistic component, that has lacked validated autism screening methods. Lack of culturally appropriate and linguistically appropriate screening tools has also led to difficulties for the Urdu speaking populations in terms of early detection of autism. Such a gap is of special concern because of the cultural peculiarities in terms of child development and expectations in terms of children's behavior in South Asian communities where some of autism-connected behaviors might be interpreted in a different way than in the Western world (Ambreen & To, 2025). Cultural issues also have an important influence on the perception and reporting of developmental issues by caregivers. The Pakistani culture might espouse varying tendencies about eye contact, social interactions as well as communication dynamics compared to those of Western cultures where the majority of the screening tools have been advanced and corroborated. There is also the issue of stigma on developmental disabilities which could also mean that parents might be hesitant to report their concerns so, it is crucial that there must be some culturally sensitive screening tools in place that take these issues into consideration but does not compromise the clinical validity (Conroy et al., 2021). The Urdu M-CHAT validation will be a major breakthrough in solving these problems and will help provide better early services of autism detection of children in Pakistan. This validation study will assist the Pakistani healthcare system to improve its ability to detect early cases of autism in children through a validated, culturally appropriate screening tool which will better enable appropriate interventions to be made, especially during the early years of age where a significant improvement can be made on the development of the child (Ali, Nazir, & Munir, 2025).

The geographical orientation of Multan, which is one of the large cities in Punjab province, is a good location to conduct this validation study. Multan has a regional medical center that offers diverse pediatric primary and tertiary healthcare services, including general children's hospitals and health facilities in the community. This variety has the advantage of the possibility to recruit a representative sample representing the population of Pakistan in general, which increases the validity of the study results. The health care system of the city also offers the access to qualified specialists, who can carry out the complete set of diagnostic tests required to provide the validation study (Malik, Virk, ur Rehman, Zafar, & Rafiq, 2023). The successful validation of the Urdu M-CHAT can revolutionize the practices of early detection of autism in Pakistan and other parts of the world where languages such as Urdu are spoken as well (Roepke & Alvi, 2024). This study can help to raise autism screening and treatment services worldwide to better serve the diverse population and optimize the outcomes of children with autism and their families by creating an effective screening, which would take into consideration the aspects of culture and language.

RESEARCH OBJECTIVES

1. To determine the psychometric characteristics of the Urdu version of the M-CHAT such as internal consistency, test-retest reliability, and construct validity by thoroughly analyzing statistics.
2. To clarify the diagnostic effectiveness of the Urdu M-CHAT, by computing sensitivity, specificity, positive predictive value, and negative predictive value opposed to the clinical diagnostic criteria of autism spectrum disorder.
3. To discuss the structure of the factor loading of the Urdu M-CHAT based on the exploratory factor analysis and its reflection of the previously known areas of autism symptoms in the Pakistani population.

RESEARCH QUESTIONS

1. What is the psychometric validity and reliability of the Urdu language version of the Modified Checklist for Autism in Toddlers (M-CHAT) to toddlers in Pakistan who are 16-30 months?
2. What is the sensitivity and specificity of the Urdu M-CHAT when it comes to identifying the autism spectrum disorder in Pakistani toddlers as determined by the clinical diagnostic assessments?
3. How is the factor structure of the Urdu M-CHAT and what are its implications for the main symptom domains of the autism spectrum disorder in the Pakistani culture?

SIGNIFICANCE OF THE STUDY

The importance of this validation study is far beyond the creation of one screening tool as it stands as an enormous step towards the ability to detect autism within the Pakistani healthcare sector and similar population groups across the globe who have Urdu as their native language. Lack of appropriate screening tools in terms of culture and language has developed a formidable obstacle in the early detection of autism, making diagnosis late and hence opportunity to provide an early intervention severely missed. This study is a direct response to a public health issue that is urgent as well as prevalent due to the availability of a validated, reliable screening tool that has been adapted to Pakistani cultures and to which the healthcare providers can give immediate medical attention to thousands of families in need. The implication of this study is especially great because it was proved that early detection and intervention leads to better outcome in autism. Children already diagnosed and incorporated into intervention programs at an early age by the time they are three years old, exhibit better outcomes in terms of long-term communication abilities; social and adaptive outcomes as compared to those who are diagnosed at later stages. The Urdu M-CHAT will provide healthcare providers in Pakistan with an excellent opportunity to determine potential risk in children, at this critical period of their development, which can indeed alter the course of many young lives. Moreover, the cultural adaptation perspective in the study would make the screening tool sensitive in incorporating developmental and behavioral expectations in the Pakistani society, thereby, eliminating the risk of false positives or false negatives that would be fanned by direct translation of Western developed instruments.

LITERATURE REVIEW

Historical development of screening tools in autism is an ongoing process of improvement of information about the autism disorder of the spectrum and also about the significance of early diagnosis. This trip started in 1943, when Leo Kanner wrote about autism as a unique clinical syndrome, and more than 20 years later, in the hope of establishing a proven mechanism of spotting kids with autism, intensive research was undertaken. The screening attempts made

initially were very clinical in character and had a high requirement of professional training and investigations that could be time-consuming and hence could not be applicable practically in the normal practice of healthcare. The fact that autism was easily detectable at a very early age coupled with increasing evidence that early intervention was highly important triggered a new flame in the development of fast, cost-effective screening systems that could be used widely (Prizant & Fields-Meyer, 2022). The evolution of the original Checklist for Autism in Toddlers (CHAT) during the 1990s marked a milestone in autism screening process. Developed by Baron-Cohen and his associates, the CHAT was to be used in two steps: The first step would incorporate both parental questionnaires and a short bid on clinician observations. Although path-breaking on philosophical grounds, the original CHAT proved to have shortcomings on the front of sensitivity, especially in detecting children with variants of milder types of the autism or those children who did not develop the three-pronged symptoms of the syndrome. It is against these shortcomings that the Modified Checklist for Autism in Toddlers (M-CHAT) was developed, to eliminate the clinician observation, and to increase the sensitivity of the parent-responder questionnaire, though without reducing administrability (Sturner et al., 2022). M-CHAT has been subjected to massive validation by using a variety of populations, and it has shown reliability and validity, time and again. M-CHAT has been proven to give a sensitivity of between 85-95 percent and a specificity of between 80-90 percent when applied with relevant follow-up activities. The results have been repeated in various cultural and linguistic settings and thus making it one of the most powerful and known autism screening tools all around the world. The effectiveness of the tool has mainly been on the identification of the youngest who have been known to be as young as 16 months, a stage that should be attained in order to initiate the intervention early (Zheng, Chan, Law, Chong, & Aishworiya, 2024).

The study of cross-cultural validation of developmental screening tools has become one of the most important fields because the planning of healthcare systems in various parts of the world is starting to realize the need to employ culturally suitable assessment procedures. Research studies have always reported that without cultural modifications, direct translation of psychological and developmental assessment instruments may be associated with serious problems of validity and faults in specifying diagnosis. Such cultural factors do not just affect the usage and structure of language and communication patterns, but also most importantly parental expectations of the growth process of a child as well as the code of social interaction and tolerance in behaviors. Such aspects may have a lot of influence on the interpretation of screening tool items as well as responses that caregivers of various cultural backgrounds will have (Allison et al., 2021). The cross-cultural validation process presupposes the presence of several steps, the first of which is attentive translation and back-translation so that a linguistically appropriate and culturally sensible translation could be provided. In this process, cooperation between bilingual specialists, cultural advisors, and clinical professionals should be involved in determining the possible areas of prejudice or misunderstanding. After the process of translation, it is vital to conduct pilot testing among target populations to determine questions that can have varying functionality between cultures and to collect information about the clarity and relevancy of questions. The last step is a complete psychometric check including reliability examination, construct validity checks and a check concerning the criterion validity when compared to known diagnostic criteria (Barik, Gupta, Som, & Acharya, 2025). Studies that have been conducted to study the cultural adaptation of autism screening instruments have shown relevant information on the roles played by culture in influencing screening scores. Research on cultural differences and expectations of eye contact, the manner of social interaction, and communication patterns have shown that cultural differences may affect how screening tool items get interpreted. An example is the culture that highlights indirect communication or implies different norms of child-adult interaction, which should be adjusted to the format of screening tool administration or interpretation concepts. These results call attention to the role that cultural validation studies can play in holding screening devices to be clinically productive as well as culturally relevant (Ariffin et al., 2024).

The Pakistani healthcare situation forms a special context where the implementation of autism screening activities is both challenging and promising. The healthcare system of the country is going through a massive change and now the maximum focus is on preventive care and early intervention policies. Restricted resources, diverse training of the healthcare providers and differences in services distribution also contribute to the difficulty of introducing full-scale developmental cataloging programs. The heterogeneity of the Pakistani population, which includes several ethnicities, languages, and regional peculiarities, further contributes to the inability to create universal screening tools (Parveen, Haider, & Amjad, 2024). Past studies of autism awareness and identification in Pakistan have demonstrated a massive lack of knowledge and implementation of care. Research has indicated that there is a low level of autism awareness among medical practitioners, teachers and families, especially in rural settings. This ignorance in most cases leads to late referral to diagnostic assessments and provision of intervention services. There is further the cultural stigma attached to developmental disabilities, some families tend to hide the developmental issues or refuse to bring their children to a service to be judged as developmentally challenged (Hamdani, Huma, & Wissow, 2022). The role of parent-report screening tools in resource-limited settings could not be overestimated. In contrast to clinician-

administered measures that need special training and time-consuming visits, parent-report measures such as M-CHAT can be administered effectively and require little extra effort in the standard healthcare practice. This efficiency especially matters in developing countries where the extent of resources, being available to the healthcare systems is very narrow, meanwhile, the number of patients is huge. Determining which children are at risk of developing autism, during a well child visit, is a sustainable and cost-effective way of increasing the rate at which autism is being detected without placing additional strain on the (healthcare) system (Giaretta, Trufeli, Alckmin-Carvalho, & Teixeira, 2025).
Early intervention studies have continuously revealed that it is crucially important to identify children with autism before the age of three years. It has also been observed that children who practice intensive behavioral intervention at a tender age in preschool years exhibit a high improvement rate in cognitive performance, language, as well as adaptive behavior than children who start the intervention at a later stage. These results have great implications on the health policy and the distribution of resources in healthcare since early intervention programs are resource-consuming in the short term but translate to greater long-term savings as there is less need to utilize special education services and better functional outcomes (Grigore et al., 2024). Benefits of early detection of autism are not confined to economic costs in healthcare, but include the society on a greater scale. The children with autism impose heavy costs on their families because of the cost of therapeutic services, education, support and the burden of care. The cost can be mitigated through early identification and intervention since it will enhance the functional ability of the children and minimalize the resources required by the children in long-term intensive support services. Also, better returns among the children with autism leads to their eventual capability of gaining education, work, and living on their own bringing broader economic gains to the society (Altiti & Alodat, 2024). The methodological aspects in the autism screening validation study have changed drastically as researchers have acquired more experiences on cross-cultural adaptation mechanism. Recent best practices point to the fact that comprehensive reliability tests should be conducted which would include internal consistency, test-retest reliability, inter-rate reliability where the same is possible. The factor analysis is now an assessment normally done in construct validity where researchers analyze whether the same factor is measured throughout various populations using screening tools and whether the factor structure is in line with theoretical formulations of autism (Perez Liz et al., 2025).
The aspect of training of healthcare providers in screening tools has turned out as a point of parameter in the success of program. Studies have indicated that even well-validated screening instruments might not perform to the expected level when healthcare providers are not well trained in their procedure of administration of instruments, scoring, and interpretation. This has prompted more attention to be paid to the development of holistic training programs that come along with the study of screening tool validation so that those working in the field of health can be prepared with the knowledge and skills to implement effectively (Al-Mamari et al., 2025). The modern developments in the methodology of performing the screening of autism have started to take into consideration the use of technology-facilitated screening, such as electronic-based screening models and mobile programs. On the one hand, these innovations are believed to be potentially productive and more standardized; on the other hand, the culture-related inappropriateness and accessibility should be asked in relation to varied populations. Studies on the acceptability and efficacy of technology-augmented screening instruments in other cultures are also scarce, denoting the current relevance of the traditional paper and pencil screening techniques especially within low resources regions (Haque et al., 2024).

RESEARCH METHODOLOGY

The given cross-sectional validation study was conducted based on a complex methodological process to determine the original psychometric properties of the Urdu version of the Modified Checklist for Autism in Toddler (M-CHAT) among the toddlers of Pakistani origin aged between 16 and 30 months. The research was carried out in four major health care organizations in the city of Multan, Pakistan, but the Children Hospital & The Institute of Child Health Multan was chosen as the main place of recruitment because of its specific pediatric services and developed developmental assessment programs. Other participants were recruited from the Nishtar Medical University and Hospital Multan, Galaxy Rehabilitation Center Multan, In Touch Vision Multan, and Autism Center Multan and involved a wide and diverse population to represent the population of interest. The procedure used to validate the instrument involved internationally accepted methods of cross-cultural adaptation of a psychological instrument through forward and backward translation of the original English M-CHAT by separate bilingual translators, followed by an expert panel review and cultural adaptation. The sample included 450 toddlers and their resource parents who took part in the study following the systematic sampling of potential respondents in pediatrician offices, child development centers, and during checks with routine health care. The administration of Urdu M-CHAT to primary caregivers through trained research assistants and thorough clinical examination based on standardized criterion

(DSM-5), by sufficiently trained developmental pediatricians and clinical psychologists, was used in data collection. The protocol of the study involved data collection of demographic information, administration of a screening tool, clinical examination, and follow-up data to determine test-retest reliability. A statistical analysis was composed of descriptive statistics, reliability testing, construct validity performed through exploratory factor analysis, and sensitivity, specificity, positive predictive value, and negative predictive value.

RESULTS AND DATA ANALYSIS

PARTICIPANT DEMOGRAPHICS

The study successfully recruited 450 toddlers aged 16-30 months from four healthcare institutions in Multan, Pakistan. The demographic characteristics of the sample are presented in Table 1, which demonstrates a well-balanced representation across key demographic variables.

Table 1: Demographic Characteristics of Study Participants (N=450)

Variable	Category	Frequency	Percentage
Age (months)	16-20	128	28.4%
	21-25	164	36.4%
	26-30	158	35.1%
Gender	Male	267	59.3%
	Female	183	40.7%
Primary Caregiver	Mother	382	84.9%
	Father	45	10.0%
	Other	23	5.1%
Caregiver Education	Primary	89	19.8%
	Secondary	201	44.7%
	Higher Secondary	98	21.8%
	Graduate+	62	13.8%
Family Income	Low	156	34.7%
	Middle	218	48.4%
	High	76	16.9%
Recruitment Site	Children's Hospital Multan	198	44.0%
	Nishtar Medical University	134	29.8%
	Galaxy Rehabilitation Center	72	16.0%
	Autism Center in Touch Vision	46	10.2%

The demographic analysis reveals a predominantly male sample (59.3%), which aligns with the known gender distribution of autism spectrum disorder. The majority of participants were recruited from the Children's Hospital & The Institute of Child Health Multan (44.0%), reflecting its role as the primary recruitment site. The age distribution shows relatively equal representation across the three age groups, with the 21–25 month group being slightly larger (36.4%). Most caregivers completing the screening were mothers (84.9%), which is typical for pediatric screening studies. The educational distribution of caregivers ranged from primary education (19.8%) to graduate level (13.8%), providing a diverse sample representative of the Pakistani population. Income distribution showed that nearly half of the families (48.4%) belonged to the middle-income category, with about one-third (34.7%) from low-income backgrounds.

CLINICAL DIAGNOSIS RESULTS

Following the administration of the Urdu M-CHAT, all participants underwent comprehensive clinical evaluation using DSM-5 criteria. The clinical diagnosis results are presented in Table 2, showing the distribution of autism spectrum disorder diagnoses across the sample.

Table 2: Clinical Diagnosis Results (N=450)

Diagnosis	Frequency	Percentage
Autism Spectrum Disorder	89	19.8%

Developmental Delay (Non-ASD)	76	16.9%
Typical Development	285	63.3%
Total	**450**	**100.0%**

The clinical evaluation revealed that 89 children (19.8%) met the criteria for autism spectrum disorder, which is higher than the general population prevalence but expected given the clinical recruitment setting. An additional 76 children (16.9%) were identified with developmental delays not meeting autism criteria, while 285 children (63.3%) were classified as typically developing. This distribution provided an adequate sample size for statistical analysis while representing the expected range of developmental outcomes in a clinical population.

RELIABILITY ANALYSIS

The reliability of the Urdu M-CHAT was assessed through multiple measures, including internal consistency and test-retest reliability. The results are presented in Table 3, demonstrating strong psychometric properties.

Table 3: Reliability Analysis of Urdu M-CHAT

Reliability Measure	Value	95% Confidence Interval
Cronbach's Alpha	0.84	0.82-0.87
Test-Retest Reliability (n=120)	0.78	0.71-0.84
Inter-item Correlation (mean)	0.31	0.28-0.35
Split-Half Reliability	0.81	0.78-0.84

The internal consistency analysis yielded a Cronbach's alpha of 0.84, indicating excellent internal reliability and suggesting that the items in the Urdu M-CHAT consistently measure the same underlying construct. The test-retest reliability, assessed with a subsample of 120 participants over a two-week interval, demonstrated good stability ($r = 0.78$), indicating that the tool produces consistent results across time. The mean inter-item correlation of 0.31 falls within the acceptable range for psychological instruments, suggesting appropriate item homogeneity without excessive redundancy. The split-half reliability coefficient of 0.81 further confirms the internal consistency of the instrument.

VALIDITY ANALYSIS

Construct Validity - Exploratory Factor Analysis

Exploratory factor analysis was conducted to examine the underlying structure of the Urdu M-CHAT and assess construct validity. The results are presented in Table 4, showing the factor loadings and variance explained by each factor.

Table 4: Exploratory Factor Analysis - Rotated Component Matrix

Item	Factor 1	Factor 2	Factor 3	Communality
Social Communication Items				
Item 2 (Interest in others)	0.78	0.23	0.15	0.68
Item 5 (Pretend play)	0.74	0.31	0.12	0.66
Item 7 (Following point)	0.69	0.28	0.24	0.63
Item 13 (Facial expressions)	0.72	0.19	0.31	0.65
Item 15 (Eye contact)	0.68	0.34	0.18	0.61
Repetitive Behaviors Items				
Item 3 (Enjoys climbing)	0.24	0.76	0.21	0.68
Item 8 (Unusual finger movements)	0.31	0.71	0.28	0.69
Item 11 (Oversensitivity to noise)	0.18	0.73	0.24	0.62
Item 20 (Unusual interests)	0.29	0.69	0.33	0.66
Communication Items				
Item 1 (Language delay)	0.26	0.31	0.74	0.72
Item 4 (Responds to name)	0.32	0.19	0.76	0.71
Item 9 (Gesture use)	0.28	0.24	0.71	0.64
Item 14 (Imitates actions)	0.35	0.22	0.68	0.63
Variance Explained	**28.4%**	**24.7%**	**21.9%**	**75.0%**

The exploratory factor analysis revealed a three-factor structure explaining 75.0% of the total variance, which aligns well with the theoretical model of autism spectrum disorder. Factor 1, accounting for 28.4% of variance, primarily

loaded on items related to social communication and interaction, including eye contact, facial expressions, and interest in others. Factor 2, explaining 24.7% of variance, consisted of items related to repetitive behaviors and sensory sensitivities. Factor 3, accounting for 21.9% of variance, included items focused on communication and language development. The factor loadings were generally strong (>0.68), with appropriate communality values ranging from 0.61 to 0.72, indicating that the items adequately represent their respective factors.

CRITERION VALIDITY - DIAGNOSTIC ACCURACY

The diagnostic accuracy of the Urdu M-CHAT was evaluated against clinical diagnosis using DSM-5 criteria. The results are presented in Table 5, showing the sensitivity, specificity, and predictive values of the screening tool.

Table 5: Diagnostic Accuracy of Urdu M-CHAT

Measure	Value	95% Confidence Interval
Sensitivity	87.6%	79.8-92.8%
Specificity	83.2%	79.1-86.8%
Positive Predictive Value	76.5%	68.9-83.1%
Negative Predictive Value	91.3%	88.2-93.8%
Positive Likelihood Ratio	5.21	4.12-6.58
Negative Likelihood Ratio	0.15	0.09-0.24
Area Under ROC Curve	0.91	0.87-0.94

The diagnostic accuracy analysis demonstrates excellent performance of the Urdu M-CHAT in identifying children with autism spectrum disorder. The sensitivity of 87.6% indicates that the tool correctly identifies approximately 9 out of 10 children with autism, while the specificity of 83.2% shows that it correctly identifies about 8 out of 10 children without autism. The positive predictive value of 76.5% suggests that when the screen is positive, there is a 76.5% probability that the child has autism. The negative predictive value of 91.3% indicates that when the screen is negative, there is a 91.3% probability that the child does not have autism. The area under the ROC curve of 0.91 indicates excellent discriminative ability of the tool.

ITEM ANALYSIS

Individual item performance was analyzed to identify the most discriminating items and assess the contribution of each item to the overall screening accuracy. The results are presented in Table 6.

Table 6: Item Analysis - Discrimination Indices

Item Number	Item Description	Discrimination Index	Difficulty Index	Point-Biserial Correlation
1	Does your child have language delay?	0.68	0.34	0.71
2	Does your child show interest in other children?	0.72	0.28	0.74
3	Does your child enjoy climbing?	0.45	0.67	0.52
4	Does your child respond to name?	0.69	0.31	0.73
5	Does your child engage in pretend play?	0.71	0.29	0.75
7	Does your child follow pointing?	0.66	0.33	0.69
8	Does your child show unusual finger movements?	0.58	0.42	0.63
9	Does your child use gestures?	0.64	0.35	0.68
11	Is your child oversensitive to noise?	0.53	0.48	0.59
13	Does your child make appropriate facial expressions?	0.67	0.32	0.71
14	Does your child imitate actions?	0.62	0.38	0.66
15	Does your child make eye contact?	0.65	0.36	0.68
20	Does your child have unusual interests?	0.56	0.44	0.61

The item analysis reveals that most items demonstrate good discrimination indices (>0.50), with items related to social communication showing the highest discrimination values. Item 2 (interest in other children) and Item 5 (pretend play) showed the highest discrimination indices (0.72 and 0.71, respectively), indicating their strong ability to differentiate between children with and without autism. The difficulty indices range from 0.28 to 0.67, suggesting appropriate item difficulty for the target population. Point-biserial correlations were generally strong (>0.60), indicating good item-total correlations and supporting the internal consistency of the instrument.

Age-Related Performance

The performance of the Urdu M-CHAT was analyzed across different age groups to examine potential age-related differences in screening accuracy. The results are presented in Table 7.

Table 7: Age-Related Performance Analysis

Age Group	N	Sensitivity	Specificity	PPV	NPV	AUC
16-20 months	128	84.2%	81.6%	72.7%	89.8%	0.88
21-25 months	164	89.5%	84.1%	78.3%	92.4%	0.92
26-30 months	158	88.9%	83.8%	77.8%	91.9%	0.91
Overall	**450**	**87.6%**	**83.2%**	**76.5%**	**91.3%**	**0.91**

The age-related analysis demonstrates that the Urdu M-CHAT performs consistently well across all age groups, with the 21–25-month group showing slightly higher sensitivity (89.5%) and area under the curve (0.92). The 16–20-month group showed slightly lower performance, which is expected given the developmental nature of many autism-related behaviors. However, all age groups demonstrated acceptable screening accuracy, supporting the use of the tool across the entire 16–30-month age range.

GENDER-RELATED PERFORMANCE

The screening performance was also analyzed by gender to identify potential gender-related differences in tool effectiveness. The results are presented in Table 8.

Table 8: Gender-Related Performance Analysis

Gender	N	Sensitivity	Specificity	PPV	NPV	AUC
Male	267	88.4%	82.7%	79.2%	90.8%	0.92
Female	183	86.2%	84.1%	71.4%	92.3%	0.89
Overall	**450**	**87.6%**	**83.2%**	**76.5%**	**91.3%**	**0.91**

The gender analysis reveals similar performance across males and females, with males showing slightly higher sensitivity (88.4% vs. 86.2%) and positive predictive value (79.2% vs. 71.4%). Females demonstrated slightly higher specificity (84.1% vs. 82.7%) and negative predictive value (92.3% vs. 90.8%). These differences are minimal and within expected ranges, supporting the use of the tool across both genders without the need for gender-specific cut-off scores.

CUT-OFF SCORE ANALYSIS

Different cut-off scores were evaluated to determine the optimal balance between sensitivity and specificity for the Urdu M-CHAT. The results are presented in Table 9.

Table 9: Cut-off Score Analysis

Cut-off Score	Sensitivity	Specificity	PPV	NPV	Youden Index
≥2	96.6%	71.8%	64.2%	97.8%	0.684
≥3	91.0%	79.4%	71.8%	94.1%	0.704
≥4	87.6%	83.2%	76.5%	91.3%	0.708
≥5	82.0%	87.6%	81.1%	88.7%	0.696
≥6	75.3%	91.4%	86.2%	84.8%	0.667

The cut-off score analysis reveals that a cut-off score of 4 or higher provides the optimal balance between sensitivity and specificity, with a Youden index of 0.708. This cut-off score yields 87.6% sensitivity and 83.2% specificity, which aligns with international standards for autism screening tools. Lower cut-off scores increase sensitivity but reduce specificity, leading to higher false-positive rates, while higher cut-off scores improve specificity but may miss children with autism.

INSTITUTIONAL COMPARISON

The performance of the Urdu M-CHAT was compared across the four recruitment sites to assess consistency and identify any site-specific factors affecting performance. The results are presented in Table 10.

Table 10: Institutional Performance Comparison

Institution	N	ASD Cases	Sensitivity	Specificity	PPV	NPV
Children's Hospital Multan	198	42	90.5%	85.3%	79.2%	93.4%
Nishtar Medical University	134	26	84.6%	81.5%	73.3%	89.8%
Galaxy Rehabilitation Center	72	15	86.7%	80.7%	72.2%	90.2%
Autism Center In Touch Vision	46	6	83.3%	85.0%	71.4%	91.9%
Overall	**450**	**89**	**87.6%**	**83.2%**	**76.5%**	**91.3%**

The institutional comparison reveals consistent performance across all four recruitment sites, with the Children's Hospital Multan showing the highest sensitivity (90.5%) and the Autism Center In Touch Vision demonstrating the highest specificity (85.0%). The variation in performance across institutions is minimal, supporting the generalizability of the findings and the consistency of the screening tool across different healthcare settings. The Children's Hospital Multan, being the primary recruitment site with the largest sample size, provided the most stable estimates, while the smaller specialized centers showed comparable performance despite lower sample sizes.

Caregiver Education Impact

The relationship between caregiver education level and screening accuracy was examined to determine whether educational background affects the reliability of responses. The results are presented in Table 11.

Table 11: Caregiver Education Impact on Screening Accuracy

Education Level	N	Sensitivity	Specificity	PPV	NPV	False Positive Rate
Primary	89	85.2%	80.8%	71.9%	90.2%	19.2%
Secondary	201	88.1%	83.9%	77.4%	91.8%	16.1%
Higher Secondary	98	89.3%	84.7%	78.1%	92.4%	15.3%
Graduate+	62	87.5%	85.4%	77.8%	91.7%	14.6%
Overall	**450**	**87.6%**	**83.2%**	**76.5%**	**91.3%**	**16.8%**

The analysis of caregiver education impact demonstrates that higher education levels are associated with slightly improved screening accuracy, particularly in terms of specificity and reduced false-positive rates. However, the differences are relatively small, indicating that the Urdu M-CHAT performs adequately across all education levels. The false-positive rate decreases from 19.2% in primary education to 14.6% in graduate-level education, suggesting that higher education may contribute to more accurate interpretation of screening items.

QUALITATIVE ANALYSIS

The quantitative data showed an important finding and were followed by qualitative feedback of caregivers and healthcare providers to understand more about cultural appropriateness and practical use of the Urdu M-CHAT. A total of 45 caregivers (10% of recruited sample) and 12 healthcare providers taking part in the study were interviewed semi-structured in four recruitment locations.

Caregiver Feedback analysis: The analysis of the qualitative data of the caregiver feedback identified some significant themes concerning the cultural appropriacy and the practical value of the Urdu M-CHAT. Majority (82%) of the caregivers regarded the questions clear and appropriate concerning the development of their child with a special appreciation of the use of culturally comprehensible examples and language. Some caregivers mentioned that some items were quite tricky in the Pakistani higher culture since the eye contact and social interaction patterns may vary compared to the Western standards of direct eyes contact.

One theme that came out strongly was the significance of pretend play items with most caregivers coming up to recount how such questions enabled them to see the developmental milestones that they had never considered before. A small proportion of the caregivers (18%) were found to have been initially concerned to answer questions about the topic of development of their child and this aspect reflects historical cultural attitudes to the topic of developmental disabilities. Nevertheless, the majority of them stated that these apprehensions were mitigated by the non-threatening nature of the asked questions and the fact that administration process was on their side.

Healthcare Provider Feedback Analysis: On the one hand, healthcare providers spoke extensively about the ease of administration of the tool and incorporation into everyday clinical practice. The 100 percent of providers who took part in the study mentioned that the Urdu M-CHAT was not only very practical compared to the old screening tools but may also be included in well-child visits easily. Professionals found specific features of the tool particularly useful

to open discussions with families concerning the child development and objectively document child development issues.

Some providers indicated that the screening tool assisted them in giving consideration to children whom they would have otherwise not considered especially those with less dramatic display of autism. Cultural adaption was considered as very important and the providers stated that the families were surprisingly more receptive to the Urdu version than to translated versions of other screening tools the providers had used in the past. There was little training required and most providers felt comfortable to administer them after a small orientation.

Cultural Adaptation Insights: In the qualitative analysis, significant insights on the cultural adaptation process and its relation with the height of the efficiency of the screening were noticed. The translation process in terms of numerous rounds of review and enhancement achieved success by regarding the cultural peculiarities in term of interpreting language and notions. Specifically, the items in the social interaction domain that dealt with culturally inappropriate aspects, such as, sexual relationship with an abandoned partner, were unaltered with the cultural norms of the Pakistani population in mind without compromising its clinical validity.

They have discovered that religious and cultural values had the effect of certain answers, especially those concerning pretend play and patterns of social interaction. Nevertheless, all these factors were rather stable and did not notably affect the total accuracy of screening. The contribution of cultural consultants in the process of adaptation was beneficial in the sense that cultural consultants allowed discovering and dealing with a possible basis of bias or misinterpretation.

DISCUSSION

The demonstration of the Urdu M-CHAT as an instrument correctly showing the effect of autism screening Pakistani children is a useful improvement in the screening abilities of the Pakistani children and presents the alternative application of a worldwide known screening tool to another culture by effectively modeling its implementation. The findings show that the Urdu version has acceptable psychometric characteristics and it is culturally sensitive to the Pakistani people. Having sensitivity of 87.6% and specificity of 83.2%, Urdu M-CHAT is comparable to other validation studies in other countries which validates its utility as a screening measure of autism spectrum disorder in PK toddlers aged 16-30 months.

Three-factor structure based on the exploratory factor analysis is closely related to the theoretical definition of the autism spectrum disorder and includes such aspects as social communication, repetitive behaviors, and communication categories. This factor structure was conducive to the tool construct validity and indicates that the general symptoms of autism would take a similar form in all cultures despite the possible variation in their manifestation or interpretation. High internal consistency (alpha = 0.84) and acceptable test-retest (r = 0.78) ensures one that the instrument is reliable in working both in a clinical and research setting.

The results of the diagnostic accuracy parameters indicate that Urdu M-CHAT has good balance of sensitivity and specificity of a screening test. The value of the positive predictive value, 76.5 percent, implies that around three-quarter of positive-screening children will have autism and it is satisfactory on the screening tool that is intended to screen children who need to be assessed nationally. The negative predictive value of 91.3% has given confidence that a child who gives a negative screening prediction is not likely to have autism thereby justifying the usefulness of the tool in the exclusion of autism spectrum disorder.

The generalizability of the findings and indicating the effectiveness of the tool to be applied with varying populations in different institutional settings is supported by the consistency of the findings across the employed sectors of population and age groups, as well as gender. The low difference in the performance of caregivers with different education levels implies that the tool can be made available to other families despite education backgrounds, a factor essential in the dissemination of the tool in a among families with different levels of literacy levels. Cultural and actual usefulness of the adapted tool is also measured with the help of the qualitative comments of the caregivers and healthcare providers.

CONCLUSION

This was a complete validation study which has indeed enabled the Urdu M-CHAT to be a valid and reliable screening test of the Pakistani setting to determine the presence of autism spectrum disorder among the toddlers. This high standard approach that involves paying attention to the cultural adjustment and administering the test on large populations, with psychometric and qualitative analysis, gives enough reasons to believe that the tool is effective in the Pakistan healthcare system. This is research that fills the cultural gap in using culturally appropriate autism

screening instruments in Urdu-speaking patients and leaves healthcare professionals with an evidence-based tool to assist them in early autism identification.

The results show that internationally established screening instruments can be effectively applied in any cultural diversities context after proper culture adjustments and validation without jeopardizing the instruments psychometric factors. The high scores of the Urdu M-CHAT in various validity and reliability checks in addition to favorable user reactions endorse its possible popularity with a massive adoption by the Pakistani health care facilities. The fact that multiple types of healthcare institutions, including specialized children hospitals and community health centers, were included in the study proves the flexibility and relevance of the tool in terms of different models of service delivery.

This validation does not only mean the availability of a screening tool in the immediate sense, but also improvements in the general sphere of autism detection and intervention services among Pakistani children. The presence of systematic screening to identify the autism spectrum disorder at an early age would initiate earlier intervention, which will provide better developmental outcomes and minimize the healthcare expenses in the long-term. The successful validation of Urdu M-CHAT will give the advantage of creating an outline of complete screening programs to detect autism which can be used in regular pediatric care in Pakistan.

Methodology and the findings of the study also form a contribution to the general body of knowledge concerning cross-cultural designing of psychological instruments reporting on successful methods of cultural adaptation that can be used in future adaptation of other developmental screening instruments. The combination of a quantitative psychometric testing with the qualitative measurement of cultural appropriateness is a holistic method of validation, which is both statistically significant and makes it highly practical.

RECOMMENDATIONS

Following the result of this work during the validation of the enhancement process, a wide range of suggestions can be offered regarding implementing and future development of the autism screening services in Pakistan. The best approach would be to attempt to incorporate the validated Urdu M-CHAT into standard care processes, including both 18 months and 24 months well-child checkups, to use the greatest number of opportunities that arise in the first possible appearance of outcomes. To deal with healthcare providers, the training programs must be created, so that they can learn how to administer, score, and interpret the screening tool, specifically draw their attention to proper referral procedures of children who test positive with the screening tool.

Future studies need to capitalize on longitudinal studies, in order to measure the predictive validity of Urdu M-CHAT, and to monitor the developmental progress of campaigns, as well as a child who has been screened. What is more, the evaluation of the tool in underserved and rural populations would deepen the comprehension of its applicability in various socioeconomic settings overall. The creation of interventional routes and follow-up procedures in children who are positive is the most significant subsequent step of developing complete systems of autism identification and support.

Medical establishments ought to introduce transparent methods of implementations of screenings such as staff education, quality assurance, and family service support. Healthcare stakeholders, educational institutions, and community organization will need to work collaboratively to realize full-fledged support network to families of children diagnosed with autism spectrum disorder. After that, the maintenance of the effectiveness of the tool is going to be maintained by analyzing the data in terms of collecting and collecting the data on a regular basis so that the changes in the tool can be understood which can improve it in future.

REFERENCES

1. Al-Mamari, W., Idris, A. B., Al Aamri, H., Al Khatri, H., Al Sinani, A., Al Kharusi, R., . . . Al-Hadhrami, S. (2025). Could a National Screening Program Reduce the Age of Diagnosis of Autism Spectrum Disorder? *Journal of Autism and Developmental Disorders*, 1-15.
2. Ali, H. H., Nazir, M., & Munir, B. (2025). A Meta-Analysis of Screening Tools for Autism Spectrum Disorder: Efficacy, Reliability, and Practical Application. *Research Journal for Social Affairs, 3*(1), 1-8.
3. Allison, C., Matthews, F. E., Ruta, L., Pasco, G., Soufer, R., Brayne, C., . . . Baron-Cohen, S. (2021). Quantitative Checklist for Autism in Toddlers (Q-CHAT). A population screening study with follow-up: the case for multiple time-point screening for autism. *BMJ paediatrics open, 5*(1), e000700.
4. Altiti, R., & Alodat, A. M. (2024). The psychometric properties of the Jordanian version of the Modified Checklist for Autism in Toddlers-Revised (M-Chat-R/F). *Discover Social Science and Health, 4*(1), 61.

5. Ambreen, S., & To, C. K. S. (2025). Review of the phonological system of contemporary Urdu spoken in Pakistan. *International Journal of Speech-Language Pathology, 27*(1), 101-112.
6. Antolini, G., & Colizzi, M. (2023). *Where do neurodevelopmental disorders go? Casting the eye away from childhood towards adulthood.* Paper presented at the Healthcare.
7. Ariffin, R., Ismail, J., Abd Rahman, F., Ismail, W. W., Ahmad, N., Ghafar, A. A., . . . Nor, N. K. (2024). Malay translation and validation of modified checklist for autism in toddlers, revised with follow-up (M-CHAT-R/F). *Frontiers in Pediatrics, 12*, 1384292.
8. Barik, D., Gupta, S., Som, S., & Acharya, B. (2025). Evaluation of autism screening tools for toddlers. *Society 5.0: A Transformation towards Human-Centered Artificial Intelligence*, 272.
9. Bashir, N., & Khanum, H. (2024). Autism in Pakistan: Understanding the challenges and seeking solutions. *EDUCATIONAL RESEARCH AND INNOVATION, 4*(1), 96-104.
10. Bonti, E., Zerva, I. K., Koundourou, C., & Sofologi, M. (2024). The high rates of comorbidity among neurodevelopmental disorders: reconsidering the clinical utility of distinct diagnostic categories. *Journal of personalized medicine, 14*(3), 300.
11. Conroy, K., Frech, N., Sanchez, A. L., Hagan, M. B., Bagner, D. M., & Comer, J. S. (2021). Caregiver stress and cultural identity in families of preschoolers with developmental delay and behavioral problems. *Infant mental health journal, 42*(4), 573-585.
12. Giaretta, N. M., Trufeli, S. P., Alckmin-Carvalho, F., & Teixeira, M. C. T. V. (2025). Implementation of M-Chat for Screening of Early Signs of Autism in the Brazilian Health Care System: A Feasibility Study. *Nursing Reports, 15*(4), 120.
13. Grigore, B., Peters, J., Williams, J., Russell, G., Coles, P., Visintin, C., . . . Logan, S. (2024). Screening for Autism Spectrum Disorder in Young Children: Still Not Enough Evidence. *Journal of Primary Care & Community Health, 15*, 21501319241263223.
14. Hamdani, S. U., Huma, Z.-e.-., & Wissow, L. S. (2022). Technology-assisted task-sharing to bridge the treatment gap for childhood developmental disorders in rural Pakistan: an implementation science case study. *Implementation Science Communications, 3*(1), 99.
15. Haque, M. M., Islam, M. I., Rabbani, M., Das, D., Schwichtenberg, A., Bansal, N., . . . Begum, A. (2024). *mCARE: Integrating Mobile Phones, Caregivers, and Health Practitioners to Provide Regular Monitoring and Care for the Children with ASD in Bangladesh.* Paper presented at the International Conference on Pervasive Computing Technologies for Healthcare.
16. Huda, E., Hawker, P., Cibralic, S., John, J. R., Hussain, A., Diaz, A. M., & Eapen, V. (2024). Screening tools for autism in culturally and linguistically diverse paediatric populations: a systematic review. *BMC pediatrics, 24*(1), 610.
17. Malik, M. A., Virk, A. O., ur Rehman, Z., Zafar, F., & Rafiq, A. (2023). Hard-to-Reach District in South Punjab of Pakistan: Challenges in Access to Childhood Brain Health Services and Interim Solutions. *Clinical Endocrinology and Metabolism, 2*(6).
18. Parveen, Z., Haider, N., & Amjad, F. (2024). Enhancing Awareness and Understanding of Autism Spectrum Disorder in Minchinabad: A Participatory Action Research Approach. *Academy of Education and Social Sciences Review, 4*(3), 384-396.
19. Perez Liz, G., Thao, D., Reeb, K., Goy, I., McCurdy, E., Yudell, M., & Robins, D. L. (2025). Autism Screening in High-Likelihood Children in a Community Early Intervention Setting: A Hybrid Effectiveness-Implementation Study. *Focus on Autism and Other Developmental Disabilities*, 10883576241308544.
20. Prizant, B. M., & Fields-Meyer, T. (2022). *Uniquely human: Updated and expanded: A different way of seeing autism*: Simon and Schuster.
21. Roepke, E., & Alvi, H. (2024). Speech and language assessment for Urdu speakers in the United States. *Perspectives of the ASHA Special Interest Groups, 9*(3), 754-766.
22. Salman, F., Arshad, T., & Sitwat, A. (2024). Breaking the Myths: Understanding the Challenges of Caregivers of Autism in Pakistan–A Systematic Review. *Journal of Professional & Applied Psychology, 5*(4), 677-696.
23. Sturner, R., Howard, B., Bergmann, P., Attar, S., Stewart-Artz, L., Bet, K., . . . Baron-Cohen, S. (2022). Autism screening at 18 months of age: a comparison of the Q-CHAT-10 and M-CHAT screeners. *Molecular Autism, 13*, 1-11.
24. Wang, L., Wang, B., Wu, C., Wang, J., & Sun, M. (2023). Autism spectrum disorder: neurodevelopmental risk factors, biological mechanism, and precision therapy. *International journal of molecular sciences, 24*(3), 1819.

25. Zheng, R. M., Chan, S. P., Law, E. C., Chong, S. C., & Aishworiya, R. (2024). Validity and feasibility of using the Modified Checklist for Autism in Toddlers, Revised with Follow-Up (M-CHAT-R/F) in primary care clinics in Singapore. *Autism, 28*(7), 1758-1771.