

EXPLORING THE DIMENSIONALITY OF THE Q-CHAT-10: A LATENT CLASS ANALYSIS APPROACH

* **Zakariya M. S. MOHAMMED:**

*Department of Mathematics, College of Science, Northern Border University,
Arar, Saudi Arabia*

*Center for Scientific Research and Entrepreneurship, Northern Border University,
Arar, Saudi Arabia*

Email : zakariya.mohammed@nbu.edu.sa

Reçu le: 25/06/2025 **Accepté le :** 19/10/2025 **Publication en ligne le:** 09/12/2025

ABSTRACT:

The objective of this study is to explore the latent structure of the Q-CHAT-10 by applying latent class analysis to a sample of toddlers residing in Saudi Arabia. Several model-fit indices were used to identify the optimal number of classes, and entropy was employed to assess classification clarity (values ≥ 0.80 indicating acceptable separation).

The results revealed the existence of three latent classes corresponding to low-risk, moderate-risk, and high-risk subgroups. Class 1, the low-risk group, comprises 31% of the sample and is characterized by low probabilities of autistic traits. Class 2, the largest subgroup (44%), represents the moderate-risk group and displays moderate probabilities of autistic traits across all items. Class 3, accounting for 25% of the sample, represents the high-risk group and is marked by consistently high probabilities of autistic traits.

This study contributes to both health policy and clinical practice by identifying distinct autism risk subgroups among toddlers. Classifying children into low-, moderate-, and high-risk groups facilitates early, targeted interventions and more efficient allocation of healthcare resources, potentially reducing long-term burdens on families and health systems.

Keywords: Autism Spectrum Disorder (ASD); Q-CHAT-10; Latent Class Analysis; Toddlers; Saudi Arabia.

I. INTRODUCTION

Autism spectrum disorder (ASD), commonly referred to as autism, is a developmental disability characterized by difficulties in social interaction and communication, along with repetitive behaviors [1]. In 2021, an estimated 61.8 million individuals globally were affected by autism, with a 95% confidence interval ranging from 52.1 to 72.7 million. The prevalence was approximately 1 in 127 to 1 in 100 individuals. The age-standardized prevalence rate was 788.3 per 100,000 people—1,064.7 per 100,000 among males and 508.1 per 100,000

* Zakariya M. S. MOHAMMED

among females [2,3]. Children are among the vulnerable groups with global prevalence of ASD among children is estimated at 0.51% while in Saudi Arabia, the prevalence is almost three times higher, at 1.33% [4]. Autism can be diagnosed as early as 18 to 24 months of age. Early diagnosis offers numerous benefits, including the opportunity for timely intervention, which has been shown to significantly improve developmental outcomes, facilitates access to appropriate support services, help reducing parental stress and increase the likelihood of independence in adulthood [5-8].

Several tools are available for the early screening and diagnosis of ASD, among which the Quantitative Checklist for Autism in Toddlers (Q-CHAT-10) is widely used [8–10]. Despite its widespread use, the latent structure of the Q-CHAT-10 remains understudied, particularly in non-English-speaking countries. The scoring practices often assume unidimensionality and uniform interpretation of items, which may overlook culturally specific response patterns or latent subgroups of children exhibiting similar symptoms. This gap limits the tool's diagnostic precision and hinders efforts to enhance its accuracy based on local interpretations and contextual understanding.

Latent class analysis (LCA) is a person-centered statistical methods that used to identify unobserved subgroups, or latent classes [11-13]. In the context of Q-CHAT-10, LCA offers valuable insights into the underlying heterogeneity in behavioral symptomatology among toddlers. In contrast to variable-centered approaches that assume a homogeneous population, LCA try to detect distinct response profiles that may reflect diverse behavioral subtypes. Thus, analyzing Q-CHAT-10 data can enhance understanding of its dimensionality and inform interpretation of screening outcomes.

Therefore, the objective of this study is to investigate the latent structure of the Q-CHAT-10 using latent class analysis in a sample of toddlers residing in Saudi Arabia. Specifically, the study aims to answer the following research question: What is the latent class structure underlying responses to the Q-CHAT-10 among Saudi toddlers, and what are the characteristics of the identified classes? This inquiry seeks to explore the number, nature, and interpretability of the latent classes that emerge from the data.

The remainder of this article is structured as follows: Section 2 describes the methodology, Section 3 presents the results of the LCA, and Section 4 concludes the study and highlights its limitations.

II. METHODOLOGY

Quantitative Checklist for Autism in Toddlers – 10-item (Q-CHAT-10)

The Q-CHAT-10 (Quantitative Checklist for Autism in Toddlers – 10-item version) (Figure 1) was developed by Claire Allison and colleagues at the Autism Research Centre, University of Cambridge, in 2012 [9]. It is a shortened version of the original 25-item Q-

CHAT, comprising the ten most discriminative items for screening autism in toddlers aged 18–24 months.

The scoring system is as follows:

- For items 1 to 9, assign 1 point if the response falls within columns C, D, or E; otherwise, assign 0 points.
- For item 10, assign 1 point if the response falls within columns A, B, or C; otherwise, assign 0 points.

The total score is calculated by summing the points across all 10 items. A total score of 3 or higher indicates an elevated likelihood of autism (not a diagnosis), and the child should be referred for a diagnostic assessment.

Figure 1: Q-CHAT-10 items together with degree of severity of each item

	A	B	C	D	E
A1: Does your child look at you when you call his/her name?	Always	Usually	Sometimes	Rarely	Never
A2: How easy is it for you to get eye contact with your child?	Very easy	Quite easy	Quite difficult	Very difficult	Impossible
A3: Does your child point to indicate that s/he wants something?	Many times a day	A few times a day	A few times a week	Less than once a week	Never
A4: Does your child point to share interest with you? (e.g. pointing at an interesting sight)	Many times a day	A few times a day	A few times a week	Less than once a week	Never
A5: Does your child pretend?	Many times a day	A few times a day	A few times a week	Less than once a week	Never
A6: Does your child follow where you're looking?	Always	Usually	Sometimes	Rarely	Never
A7: If someone is upset, does your child try to comfort them?	Always	Usually	Sometimes	Rarely	Never
A8: Would you describe your child's first words as:	Very typical	Quite typical	Slightly unusual	Very unusual	Doesn't speak
A9: Does your child use simple gestures?	Many times a day	A few times a day	A few times a week	Less than once a week	Never
A10: Does your child stare at nothing with no apparent purpose?	Never	Rarely	A few times a week	A few times a day	Many times a day

Latent Class Analysis

Latent class analysis (LCA) is a person-centered statistical method used to identify subgroups within a population that share common characteristics. These subgroups are referred to as latent classes [11-14]. To identify these latent classes, LCA analyzes participants' responses to categorical indicator variables. LCA equation model the probability of observing a particular response based on underlying unobserved latent classes. That is:

$$P(Y_1 = y_1, \dots, Y_K = y_K) = \sum_{c=1}^C P(C = c) \prod_{k=1}^K P(Y_k = y_k | C = c)$$

where:

- K is the number of indicator/observed variables
- C is the number of latent classes
- $y_1, \dots, y_K$ is a vector of observed indicator variables.

Several model fit indices are commonly used in latent class analysis. Among the most widely reported are the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), Sample Size Adjusted Bayesian Information Criteria (SABIC), and entropy with BIC and SABIC being the most reliable and reported indices compared to AIC [15]. Lower values of AIC and BIC indicate better model fit, while higher entropy values reflect clearer classification, with an entropy value of 0.8 or above generally considered acceptable [15].

Data and software

The study relies on publicly available data comprising 506 toddlers. The dataset is anonymized, ensuring that individual participants cannot be identified (available at: <https://www.kaggle.com/datasets/asdpredictioninsaudi/asd-screening-data-for-toddlers-in-saudi-arabia>). The sample size satisfies the recommended requirements for conducting latent class analysis [16].

Jamovi version 2.6, a free and open-source software built on the R programming language, was used to facilitate data analysis. It includes dedicated modules for conducting latent class analysis [17–20].

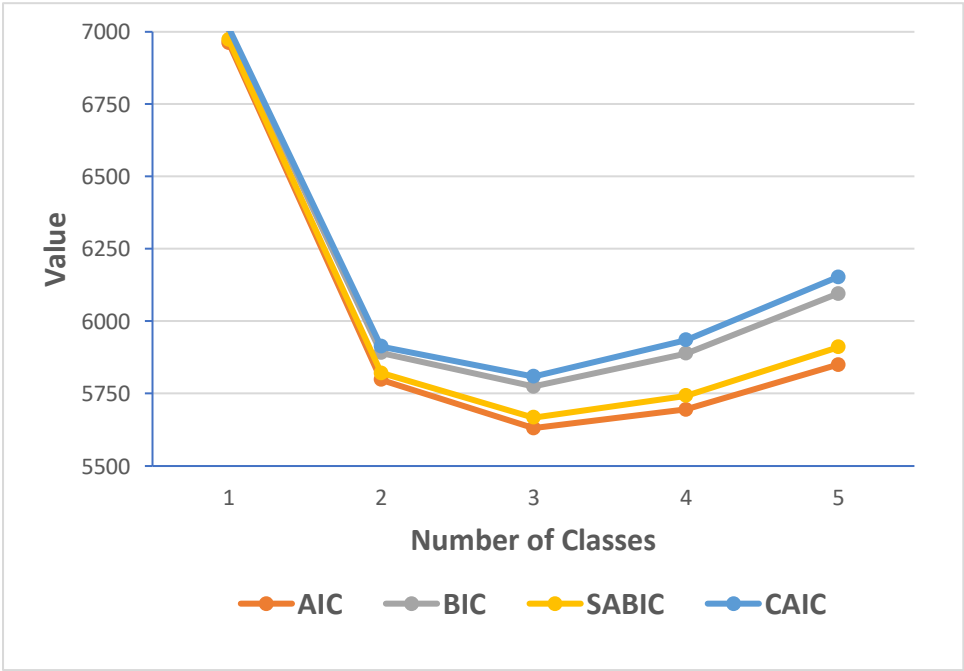
III. RESULTS

Table 1 and Figure 2 depict the model fit and classification clarity measures for latent class models with a varying number of classes ranging from one to five. All fit indices, namely, AIC, BIC, SABIC, and CAIC, show a clear improvement as the number of classes increases from one to three, with the three-class model attaining the lowest values across all criteria. Beyond three classes, the fit indices begin to increase, suggesting that the three-class model is the optimal solution. The elbow plot in Figure 1 visually supports these findings, showing a noticeable decline in fit indices up to the three-class model, followed by a slight increase. Regarding classification clarity, as measured by entropy, the two-class model achieves the highest value (0.87), but the three-class model still demonstrates acceptable classification precision (0.83). Considering both model fit indices and classification clarity, the three-class model provides the most appropriate balance and appears to be the most suitable solution.

Table 1: Model Fit Indices and Measure of Classification Clarity for varying Latent Class Solution

Number of Classes	Model Fit Indices				Classification Clarity Index
	AIC	IC	SABIC	CAIC	Entropy
1	6962	7004	6972	7014	-
2	5798	5891	5821	5913	0.87
3	5631	5775	5667	5809	0.83
4	5695	5889	5743	5935	0.85
5	5850	6095	5911	6153	0.82

Figure 2: Elbow Plot for Model Fit Indices for Latent Class Solutions with Varying Numbers of Classes



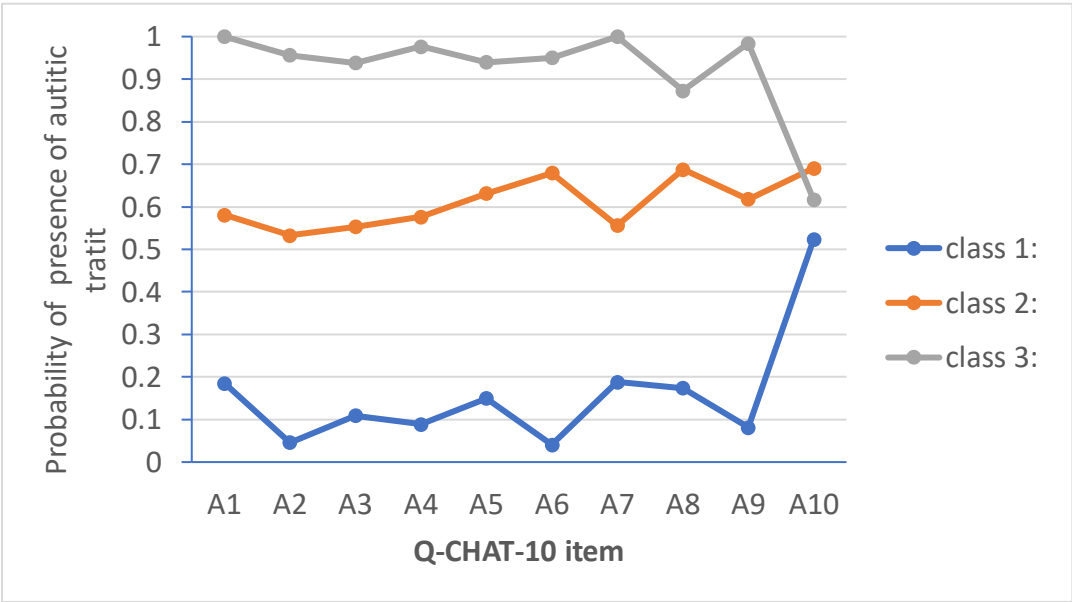
The latent class analysis identified three distinct classes of toddlers, each with its corresponding population share. These classes represent low-risk, moderate-risk, and high-risk subgroups, each with meaningful and interpretable profiles. Class 1, the low-risk group, comprises 31% of the sample and is characterized by low probabilities of autistic traits across all Q-CHAT-10 items, except for item A10, which shows a moderate probability. Class 2, the largest group at 44%, represents the moderate-risk class, exhibiting moderate

probabilities of autistic traits across all items. Class 3, accounting for 25% of the sample, is the high-risk group, characterized by consistently high probabilities of autistic traits across all items, with moderate probability for item A10 only. These findings highlight the statistical heterogeneity in the presentation of autistic traits among toddlers and support the existence of three clearly defined subgroups with varying levels of autism risk (Table 3, Figure 3).

Table 2: Detected classes

Class	Class share	Class Profile
Class 1	31%	Low-risk group characterized by low probabilities of autistic traits across all items, with a moderate probability for item A10
Class 2	44%	Moderate-risk group with average probability of autistic traits across all items.
Class 3	25%	High-risk group characterized by high probabilities of autistic traits across all items, with a moderate probability for item A10

Figure 3: Profile plot of the 3 classes



IV. CONCLUSION

The objective of this study is to examine the latent structure of the Q-CHAT-10 using latent class analysis in a sample of toddlers residing in Saudi Arabia. A total of 506 responses were analyzed using Jamovi (version 2.6). Based on multiple model fit indices—namely, AIC, BIC, SABIC, and CAIC—and the classification accuracy index (entropy), the three-class model was identified as the best-fitting solution, with an entropy value of 0.83, indicating good class separation.

The three derived classes reflect low-risk, moderate-risk, and high-risk subgroups. Class 1, the low-risk group, accounts for 31% of the sample and displays low probabilities of autistic traits across all Q-CHAT-10 items, except for item A10, which shows a moderate probability. Class 2, the largest subgroup (44%), aligns with the moderate-risk group and exhibits moderate probabilities of autistic traits across all items. Class 3, which makes up 25% of the sample, forms the high-risk group and is defined by uniformly high probabilities of autistic traits across all items, with a moderate probability observed only for item A10. These findings highlight the statistical variability in the manifestation of autistic traits among toddlers and support the existence of three distinct subgroups with varying levels of autism risk.

This study contributes to both health policy and clinical practice by identifying distinct autism risk subgroups among toddlers. The classification into low-, moderate-, and high-risk groups enables early targeted interventions and more efficient resource allocation, potentially reducing long-term burdens on families and health systems. Methodologically, the use of multiple model fit indices strengthens the robustness of the analysis. However, the study is limited by its reliance on a single screening tool and a region-specific sample, which may affect the generalizability of the findings.

Conflicts of Interest: The author declare no conflicts of interest.

Data Availability: The data that support the findings of this study are publicly available in anonymous format at (<https://www.kaggle.com/datasets/asdpredictioninsaudi/asd-screening-data-for-toddlers-in-saudi-arabia>).

V. REFERENCES

1. Shaw KA, Williams S, Patrick ME, et al. Prevalence and Early Identification of Autism Spectrum Disorder Among Children Aged 4 and 8 Years — Autism and Developmental Disabilities Monitoring Network, 16 Sites, United States, 2022. *MMWR Surveill Summ* 2025;74(No. SS-2):1–22. DOI: <http://dx.doi.org/10.15585/mmwr.ss7402a1>

2. Zeidan, J., Fombonne, E., Scorch, J., Ibrahim, A., Durkin, M. S., Saxena, S., ... & Elsabbagh, M. (2022). Global prevalence of autism: A systematic review update. *Autism research*, 15(5), 778-790. DOI: <https://doi.org/10.1002/aur.2696>
3. Santomauro, D. F., Erskine, H. E., Herrera, A. M. M., Miller, P. A., Shadid, J., Hagins, H., ... & Sankararaman, S. (2025). The global epidemiology and health burden of the autism spectrum: Findings from the Global Burden of Disease Study 2021. *The Lancet Psychiatry*, 12(2), 111-121. DOI: [https://doi.org/10.1016/S2215-0366\(24\)00363-8](https://doi.org/10.1016/S2215-0366(24)00363-8)
4. Alsulami, A. N., Alsolami, . F. N., Alsulami, . J. N. J., Alharbi, . I. M., Alotaibi, . N. F. & Alotaibi, . F. M. (2024) Prevalence and determinants of autism spectrum disorder in children in Saudi Arabia in comparison to developed countries: a systematic review and meta-analysis. *International Journal of Medicine in Developing Countries*, 8 (3), 1022-1029. DOI: <https://10.24911/IJMDC.51-1703348028>
5. Okoye, C., Obialo-Ibeawuchi, C. M., Obajeun, O. A., Sarwar, S., Tawfik, C., Waleed, M. S., ... & Mbaezue, R. N. (2023). Early diagnosis of autism spectrum disorder: a review and analysis of the risks and benefits. *Cureus*, 15(8). DOI: <https://10.7759/cureus.43226>
6. Gabbay-Dizdar, N., Ilan, M., Meiri, G., Faroy, M., Michaelovski, A., Flusser, H., Menashe, I., Koller, J., Zachor, D. A., & Dinstei, I. (2022). Early diagnosis of autism in the community is associated with marked improvement in social symptoms within 1–2 years. *Autism*, 26(6), 1353–1363. DOI: <https://doi.org/10.1177/13623613211049011>
7. Jagan, V., & Sathiyaseelan, A. (2016). Early intervention and diagnosis of autism. *Indian Journal of Health & Wellbeing*, 7(12), 1144-1148. [Available at: <https://iahrw.org/product/early-intervention-and-diagnosis-of-autism/>]
8. Elder, J. H., Kreider, C. M., Brasher, S. N., & Ansell, M. (2017). Clinical impact of early diagnosis of autism on the prognosis and parent–child relationships. *Psychology Research and Behavior Management*, 10, 283–292. <https://doi.org/10.2147/PRBM.S117499>
9. Allison, C., Auyeung, B., & Baron-Cohen, S. (2012). Toward brief “red flags” for autism screening: the short autism spectrum quotient and the short quantitative checklist in 1,000 cases and 3,000 controls. *Journal of the American Academy of Child & Adolescent Psychiatry*, 51(2), 202-212. DOI: <https://doi.org/10.1016/j.jaac.2011.11.003>
10. Sturmer, R., Howard, B., Bergmann, P., Stewart, L., Marks, D., Neer, S. M., & Williams, C. (2022). Autism screening at 18 months of age: A comparison of the Q-CHAT-10 and M-CHAT screeners. *Molecular Autism*, 13(1), 2. DOI: <https://doi.org/10.1186/s13229-021-00480-4>
11. Weller, B. E., Bowen, N. K., & Faubert, S. J. (2020). Latent class analysis: a guide to best practice. *Journal of black psychology*, 46(4), 287-311. DOI: <https://doi.org/10.1177/0095798420930932>

12. Lanza, S. T., & Cooper, B. R. (2016). Latent class analysis for developmental research. *Child Development Perspectives*, 10(1), 59-64. DOI: <https://doi.org/10.1111/cdep.12163>
13. Bauer, J. (2022). A primer to latent profile and latent class analysis. In *Methods for researching professional learning and development: Challenges, applications and empirical illustrations* (pp. 243-268). Cham: Springer International Publishing.
14. Sinha, P., Calfee, C. S., & Delucchi, K. L. (2021). Practitioner's guide to latent class analysis: methodological considerations and common pitfalls. *Critical care medicine*, 49(1), e63-e79. DOI: <https://doi.org/10.1097/CCM.0000000000004710>
15. Killian, M. O., Cimino, A. N., Weller, B. E., & Seo, C. H. (2019). A systematic review of latent variable mixture modeling research in social work journals. *Journal of Evidence-Based Social Work*, 16(2), 192–210. DOI: <https://doi.org/10.1080/23761407.2019.1577783>
16. Nylund-Gibson, K., & Choi, A. Y. (2018). Ten frequently asked questions about latent class analysis. *Translational Issues in Psychological Science*, 4(4), 440–461. DOI: <https://doi.org/10.1037/tps0000176>
17. The jamovi project. (2024). jamovi (Version 2.6) [Computer software]. <https://www.jamovi.org>
18. R Core Team. (2024). R: A language and environment for statistical computing (Version 4.4) [Computer software]. <https://cran.r-project.org> (R packages retrieved from CRAN snapshot 2024-08-07)
19. Linzer, D., & Lewis, J. (2022). poLCA: An R package for polytomous variable latent class analysis [R package]. <https://CRAN.R-project.org/package=poLCA>
20. Seol, H. (2025). snowRMM: Rasch Mixture, LCA, and Test Equating Analysis (Version 5.9.1) [jamovi module]. <https://github.com/hyunsooseol/snowRMM>