Journal of the International Neuropsychological Society (2025), 1–12
doi:10.1017/S1355617725101653

Research Article

Psychometric properties and validity of the Hong Kong version of the Penn computerized neurocognitive battery (CNB-HK) in Chinese children with and without autism spectrum disorder

Oscar Wing Ho Wong[1], Angela Man Wai Lam[2], Tyler Moore[3], Kosha Ruparel[3], Sandra Sau Man Chan[1], Raquel E. Gur[3,4] and Ruben C. Gur[3,4]

[1]Department of Psychiatry, Faculty of Medicine, The Chinese University of Hong Kong, Shatin, Hong Kong SAR, [2]Department of Psychology, Faculty of Social Science, The Chinese University of Hong Kong, Shatin, Hong Kong SAR, [3]Neurodevelopment and Psychosis Section, Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA and [4]Lifespan Brain Institute (LiBI), Children's Hospital of Philadelphia and Penn Medicine, Philadelphia, PA, USA

Abstract

Objective: Neurocognitive assessment is an essential research instrument for autism spectrum disorder (ASD), as the clinical manifestations are rooted in diverse neurocognitive processes that cause variation in clinical presentation. Few instruments comprehensively capture relevant neurocognitive domains, and most require professional assessors. The Penn Computerized Neurocognitive Battery (CNB) is widely used in child and adolescent psychiatry research across cultures. This study adapted and validated the CNB for a clinical ASD cohort in Hong Kong. **Method:** In this Hong Kong version of the CNB (CNB-HK), thirteen cognitive tasks were translated and adapted, with one task for sensorimotor speed and twelve belonging to four specific domains (episodic memory, social cognition, complex cognition, and executive function). The CNB-HK was administered to 636 normal-IQ children with ASD (mean age: 8.4 years, 87.1% male) and 412 children without ASD (mean age: 8.6 years, 55.1% male). Factor structure was examined using factor analyses. **Results:** The CNB-HK had high feasibility for children with ASD, with <7% invalid data across all tasks. The original four-factor and bi-factor structures were replicated with good model fit, and partial scalar invariance was achieved between children with and without ASD. The factor scores correlated positively with estimated IQ in the ASD group. The ASD group had worse performance across all four cognitive domains and the g factor compared to the group without ASD. **Conclusions:** The CNB-HK is a valid, multi-domain cognitive assessment tool for children with ASD in Hong Kong, offering a feasible and reliable approach for research and clinical settings.

Keywords: Neurodevelopmental disorder; Autism Spectrum Disorder; Neurocognitive assessment; Memory; Social cognition; Complex cognition; Executive function

(Received 21 May 2025; final revision 8 November 2025; accepted 13 November 2025)

Statement of Research Significance

Research Question(s) or Topic(s): Is the Hong Kong version of the Penn Computerized Neurocognitive Battery a valid and feasible tool for assessing multi-domain neurocognitive function in Hong Kong Chinese children with autism spectrum disorder?. **Main Findings**: The Hong Kong version of the Penn Computerized Neurocognitive Battery demonstrates high feasibility in assessing cognitive abilities across multiple domains in Hong Kong children with autism spectrum disorder. It enables comparisons of cognitive profiles between children with and without autism spectrum disorder. **Study Contributions**: Validation of the Hong Kong version of the Penn Computerized Neurocognitive Battery provides a culturally tailored, standardized tool for assessing neurocognitive function in Chinese children with autism spectrum disorder. This enables future research to explore cognitive mechanisms underlying autism spectrum disorder, inform targeted interventions, and facilitate cross-cultural comparisons of neurocognitive profiles, advancing both clinical practice and research in neurodevelopmental disorders.

Corresponding author: Oscar Wing Ho Wong; Email: oscarwhwong@cuhk.edu.hk
Oscar WH Wong and Angela MW Lam contributed equally.

Cite this article: Wong O.W.H, Lam A.M.W, Moore T., Ruparel K., Chan S.S.M, Gur R.E., & Gur R.C. Psychometric properties and validity of the Hong Kong version of the Penn computerized neurocognitive battery (CNB-HK) in Chinese children with and without autism spectrum disorder. *Journal of the International Neuropsychological Society*, 1–12, https://doi.org/10.1017/S1355617725101653

Introduction

Autism spectrum disorder (ASD) is a heterogeneous neurodevelopmental condition (Lombardo et al., 2019) and little is known about the drivers of its diverse presentation. As genes associated with ASD are involved in neurocognitive development (Grove et al., 2019), cognitive functioning may explain the heterogeneity of ASD. Within cognition, there is a range of domains, such as memory, language, and executive function. Notably, uneven cognitive development is well-described in ASD and executive function has prognostic implications (Joseph et al., 2002; Kenny et al., 2019). Hence, assessing the full spectrum of cognition is essential to examine its role in phenotypic variability (Bassett & Sporns, 2017; Menon & D'Esposito, 2021). While the NIH Toolbox Cognition Battery (NIH-TCB) is a feasible option (Jones et al., 2022), its psychometric properties were not examined in the ASD population. Furthermore, social cognition tests, core to ASD, are rarely included in standardized batteries.

The Penn Computerized Neurocognitive Battery (CNB) is a multi-domain cognitive assessment tool developed at the Brain Behavior Laboratory at the University of Pennsylvania. The tasks of the CNB measure accuracy, speed, and efficiency (a composite of accuracy and speed), with their construct validity established against traditional paper-and-pencil tasks (Gur et al., 2001). Its factor structure has been analyzed using correlated-factor models and bi-factor models, yielding both domain-specific factors and general factors (g) for accuracy, speed, and efficiency (Gur et al., 2010; Moore et al., 2015). For accuracy, three domain-specific factors were identified: episodic memory, social cognition, and executive/complex cognition. For speed, three factors were identified: episodic memory, deliberative (a trade-off between speed and accuracy, often termed "slow thinking"(Kahneman, 2011)), and vigilance (requiring rapid response, often termed "fast thinking"). For efficiency, four factors were identified: episodic memory, social cognition, complex cognition, and executive function. These approaches provide varied CNB performance summaries. Its automated, low-literacy design suits large-scale studies, as exemplified by the Philadelphia Neurodevelopmental Cohort (PNC) ($n = 9{,}498$; aged 8–21 years) (Calkins et al., 2015). Following the standard translation procedures, the CNB has been translated into over 25 languages, where the factor structure has been replicated in various psychiatric disorders, across the lifespan and cultures (Bhatia et al., 2017; Izgi et al., 2022; da Motta et al., 2021; Pelt et al., 2022; Scott et al., 2021; Service et al., 2020). However, the CNB has not been translated into either Mandarin or Cantonese and validated in the Chinese population. In the PNC, the CNB was successfully administered to 218 children and adolescents with ASD (aged 8–21 years), revealing that they performed worse across all neurocognitive domains compared to the group without ASD (Wong et al., 2023). In the present study, we examined the factor structure of the Hong Kong version of the CNB (CNB-HK), a Cantonese-translated version of the CNB, in children with and without ASD. We hypothesized that 1. the factor structures of the CNB-HK would be consistent with that of the original CNB, 2. CNB-HK performance correlates with estimated IQ in participants with ASD, 3. the measurement model of the CNB-HK is invariant to the ASD diagnosis, and 4. the group with ASD performs worse in the CNB compared to peers without ASD.

Method

Study population

Participants included 1,048 Chinese children with ($n = 636$) and without ($n = 412$) ASD aged 6–12 years. Children with ASD were diagnosed by child psychiatrists based on the DSM-5 criteria and recruited from a university-affiliated psychiatric center in Hong Kong (Su et al., 2024; Wan et al., 2024). Children without ASD were recruited from a territory-wide child and adolescent psychiatric epidemiological study (Chan et al., 2025), in which participants were assessed for the presence of 31 DSM-5-defined psychiatric disorders with the Diagnostic Interview Schedule for Children Version 5 (DISC-5) (Shaffer et al., 2000) and screened for ASD using the Hong Kong version of the 10-item Autism Spectrum Quotient (AQ-10-HK) (Leung et al., 2023; Wong et al., 2025). Participants who were studying in mainstream school, without history of intellectual disability, and with negative results in the DISC-5 and AQ-10-HK were recruited. Exclusion criteria for both groups included the presence of intellectual disability, neurological disorders, severe mental illness (e.g., psychosis), and major medical illnesses. Common co-occurring psychiatric conditions such as ADHD and anxiety were not considered as exclusion criteria for the ASD group. Parents or legal guardians provided written informed consent. The study protocol followed the Declaration of Helsinki and was approved by the Joint CUHK-NTEC CREC (Ref.: 2021.550).

Demographic data and IQ assessment

Sociodemographic data were obtained from caregivers. Full-scale Intelligence Quotient (FSIQ) for the children with ASD was estimated using the short form of the Hong Kong Wechsler Intelligence Scale for Children, 4th Edition [WISC-IV (HK) Short Form] (Wechsler, 2010), which included the subtests of similarities, digit span, symbol coding, and matrix reasoning.

Adaptation and pilot testing of the CNB-HK

After examination by the authors, 12 out of the original 13 tasks were deemed suitable for translation into Cantonese. The Penn Verbal Reasoning Test, a complex cognition task with verbal analogy problems, was not translated due to cultural and language specificity in the original English analogies. The Penn Conditional Exclusion Test (PCET) was originally developed as an executive function task. However, analyses of its properties revealed that it was better categorized in the domain of complex cognition (Moore et al., 2015). Therefore, we considered the PCET a task of complex cognition in the present study. To ensure there would be three tasks as indicators for each cognitive domain for the factor analyses, a Go/No-Go Task (GNG) was incorporated as an executive function task. The details of the 13 tasks are summarized in Table 1 and described elsewhere (Gur et al., 2010). We followed guidelines for the translation process (Van Widenfelt et al., 2005): The content of the 13 tasks was first translated into Cantonese by the first author, a bilingual research psychiatrist with over 10 years of experience in conducting neurocognitive examinations. Subsequently, an independent bilingual postdoctoral psychology researcher performed the back-translation. The back-translation was then compared to the original English version. The process was iterated as necessary, and any discrepancies were resolved by consensus.

https://doi.org/10.1017/S1355617725101653 Published online by Cambridge University Press

Table 1. Summary of the tasks and their respective cognitive domain in the CNB-HK

Neurocognitive domain	Cognitive function	Task
Sensorimotor speed	Motor speed	Motor Praxis Test (MP)
Episodic memory	Visual memory	Visual Object Learning Test (SVOLT)
	Word memory	Word Memory Task (CPW)
	Face memory	Face Memory Task (CPF)
Social cognition	Emotion differentiation	Emotional Differentiation Test (MEDF)
	Emotion recognition	Emotion Identification Test (ER)
	Age differentiation	Age Differentiation Test (ADT)
Complex cognition	Non-verbal reasoning	Matrix Reasoning Test (PMAT)
	Set-shifting	Penn Conditional Exclusion Test (PCET)
	Spatial ability	Line Orientation Test (VSPLOT)
Executive function	Attention	Penn Continuous Performance Test (SCPT)
	Working memory	Letter N-Back Task (SLNB)
	Inhibition	Go/No-Go Task (GNG)

Note: CNB-HK: Hong Kong version of the Penn Computerized Neurocognitive Battery.

For the Word Memory Task (CPW), the original version comprised 36 words for memorization. These words were chosen for easy comprehension to ensure that performance was not confounded by literacy levels. The translation of the 36 words adhered to two criteria: 1. Each translated word was expressed with two Chinese characters to prevent participants from relying on the number of characters for memorization, and 2. the translated Chinese word was checked against a Chinese word database by the Hong Kong Education Bureau (www.edbchinese.hk/lexlist_ch/), to ensure that the translated Chinese words were culturally common and understood by six-year-olds. Among the 36 words, all but one could be expressed using two Chinese characters. Furthermore, eight translated words were not commonly learned by six-year-olds, or not culturally common in Hong Kong. Hence, these nine words were replaced in consultation with experts from the Department of Educational Psychology of the University. The substituted words belonged to the same semantic categories as the original English words, while fulfilling the aforementioned criteria. Pilot testing was conducted with volunteers (children and adults) and the translation was refined based on feedback.

Administration, scoring of the CNB-HK and establishment of local normative data

The CNB-HK was administered by research staff and undergraduate medical students in a quiet office. Proctors received training, including sessions of supervised administration led by the first author. A standardized manual of the CNB was followed: written instructions for each task were read aloud, and understanding of the task was ensured before each task began. The 13 tasks were administered in a prefixed order, and the proctors assigned a validity score for each task on data completeness and validity. To standardize validity scoring, a written guideline outlined task-specific criteria, as invalid data vary by task. For example, an attention lapse during the registration phase of episodic memory tasks critically impacts recognition performance, whereas lapses in continuous performance tasks may reflect actual attention ability. Proctors documented scenarios potentially affecting data validity, such as participant's distractions, lack of effort, technical errors, or assessor deviations. The first author regularly reviewed these reports with proctors to ensure accurate scoring. Only complete data without significant validity concerns were used for analyses.

Except for the MP task, which is scored solely on speed, the other 12 tasks are evaluated based on both accuracy (number of correct responses) and speed (time for correct responses in milliseconds). To establish local normative data for z-transformation, we computed means and standard deviations (SD) of task accuracy and speed for each age group (age 6–12, divided by year into seven age groups, balancing males and females) among children without ASD who provided valid and complete data. After assessing the normality of the distributions, the speed scores of all tasks except for SCPT and GNG were log10-transformed before z-transformation. To reflect faster performance as higher scores, z-transformed speed scores were multiplied by −1. An efficiency score for each task was computed by summing the z-transformed accuracy and inverted speed scores, with higher scores indicating better performance. The details of the normative data are summarized in Supplementary Table S1.

Statistical analysis

Statistical analyses were performed using R version 4.4.1. Comparisons of demographics, clinical, and task-level performance of CNB-HK were examined using χ^2 tests for categorical data and independent t or Mann–Whitney's U tests for continuous data. The data structure of the CNB-HK was examined using the R package "lavaan" (Rosseel, 2012) for confirmatory factor analysis (CFA) and "psych" (Revelle, 2020) for exploratory factor analysis (EFA). CFAs were conducted using maximum likelihood estimation, and EFAs used least-squares extraction with oblimin rotation. A direct implementation of CFA was first conducted, followed by EFAs to detect subtle factor structure anomalies, such as cross-loadings specific to the study population, to aid fine-tuning the CFA model (if necessary). Although unconventional, this CFA-first approach was chosen because the factor structures of the CNB-HK were theory-driven (Moore et al., 2015). The original Moore et al. (2015) study used a CFA-first approach, as did a previous translational adaptation of the CNB (Pelt et al., 2022). Model fit indices included the root mean square error of approximation (RMSEA) (Steiger, 1990), the comparative fit index (CFI) (Bentler, 1990), and the standardized root mean square residual (SRMR): RMSEA $\leqslant 0.05$ indicates good fit ($\leqslant 0.08$ reasonable); CFI ⊠ 0.90 suggests reasonable fit (⊠0.95 preferable); and SRMR $\leqslant 0.08$ indicates good fit (Hu & Bentler, 1999). To test construct validity, correlations of the CNB-HK performance and estimated FSIQ were examined with Spearman's rho. Finally, ANCOVA and MANCOVA were used to compare CNB-HK performance between children with and without ASD. A statistical threshold of alpha $\leqslant 0.05$ was considered significant.

Results

Clinical and demographic data

The sample included 1,048 children (636 with ASD, 412 without ASD). The ASD group had more males (87.1% *vs*. 55.1%; $\chi^2 = 120$, $df = 1$, $p < 0.001$) and was slightly younger (mean ages 8.37 *vs*. 8.59; $U = 121166$, $p = 0.037$). Parental education levels were comparable between the two groups. Among the children with ASD, 42.8% had

https://doi.org/10.1017/S1355617725101653 Published online by Cambridge University Press

Table 2. Demographics of participants and comparisons between children with and without ASD

	ASD		Without ASD		Statistics		
	Mean/n	*SD/%*	*Mean/n*	*SD/%*	*t/U/χ^2*	*df*	*p*
n	636	60.69	412	39.31		–	
Age	8.37	1.71	8.59	1.69	1.21E+05	–	0.037*
Sex							
Male	554	87.11	227	55.10	120	1	<0.001***
Female	82	12.89	185	44.90			
Highest parental education attained							
Primary or less	5	0.94	3	1.88	2.25	2	0.324
Secondary	229	45.82	116	46.31			
Tertiary	397	10.13	248	11.17			
ADHD diagnosis	270	42.80	0	0		–	
IQ	97.5	16.10	N/A			–	

Note: *$p < 0.05$, ***$p < 0.001$. ASD = Autism Spectrum Disorder, ADHD = Attention-deficit/hyperactivity disorder.

Table 3. Comparisons on invalid data rate and standardized scores of the 13 CNB-HK tasks between children with and without ASD

	% Invalid		ASD		Without ASD		Statistics	
Task	ASD	Without ASD	Mean	SD	Mean	SD	*U*	*p*
MP	0	0	−0.25	1.26	0.05	0.94	114987	0.001**
SVOLT	3.60	0.20	−0.03	1.55	0.02	1.30	123514	0.627
CPW	5.72	0.82	−1.04	2.37	−0.04	1.62	36742	<0.001***
CPF	5.00	1.50	−0.07	1.58	0.09	1.46	115832	0.136
MEDF	1.90	0.50	−0.38	1.44	0.04	1.25	104657	<0.001***
ADT	2.20	0.20	0.29	2.05	0.00	1.27	125504	0.622
ER	1.60	0.00	−0.84	1.78	0.06	1.42	88793	<0.001***
PMAT	1.10	0.00	−0.01	0.90	0.00	0.66	125350	0.445
PCET	1.90	0.70	−0.21	1.64	0.09	1.40	115136	0.008**
VSPLOT	2.20	0.50	0.16	1.58	0.02	1.31	111469	0.079
SCPT	6.40	1.70	0.12	1.83	0.10	1.60	118688	0.688
SLNB	6.30	1.70	−0.23	1.79	−0.02	1.57	113493	0.119
GNG	6.10	2.40	−0.37	1.49	0.00	1.12	104145	<0.001***

Note: **$p < 0.01$, ***$p < 0.001$. All scores are z-transformed efficiency scores, except for the Motor Praxis Test, which is a z-transformed speed score multiplied by −1 to align the direction of better performance with other scores. MP = Motor Praxis Test, SVOLT = Visual Object Learning Test, CPW = Word Memory Task, CPF = Face Memory Task, MEDF = Emotional Differentiation Test, ADT = Age Differentiation Test, ER = Emotion Identification Test, PMAT = Matrix Reasoning Test, PCET = Penn Conditional Exclusion Test, VSPLOT = Line Orientation Test, SCPT = Penn Continuous Performance Test, SLNB = Letter N-Back Task, GNG = Go/No-Go Task, ASD = Autism Spectrum Disorder.

co-occurring attention-deficit/hyperactivity disorder (ADHD). The mean estimated IQ for the ASD group was 97.5 with a standard deviation of 16.1, approximating the general population's IQ distribution (Table 2).

Feasibility of CNB-HK and comparison of task-level scores

The CNB-HK was well-tolerated and could be completed by most participants in one hour. Table 3 summarizes the performance and proportion of invalid data at the task level. Children with ASD had lower sensorimotor speed (MP), performed more poorly in word memory (CPW), emotion differentiation (MEDF), emotion recognition (ER), set-shifting (PCET), and inhibition (GNG). Although the ASD group had more invalid data, the proportion was only in the range of 1.10% to 6.40% across different tasks. Most invalid data resulted from participants' inability to understand the task or lack of motivation. In the ASD group, 72.7% provided valid data for the entire CNB-HK battery (12 tasks), while 91.1% provided valid data for at least 11 out of 12 tasks. The proportion of full valid data varied by age, increasing from 53.2% for 6-year-olds to 94.1% for 12-year-olds. Among 6-year-olds with ASD, 70.3% provided valid data for 11 tasks and 80.2% for 10 tasks. The three executive function tasks (SLNB, GNG, CPT) had the highest rates of invalid data (18.4–19.3%) in this specific group (Supplementary Table S2). These missing data were handled using maximum likelihood estimation in the subsequent factor analyses.

Confirmatory factor analysis 1: correlated-factor models

The theory-driven correlated 4-factor CFA model of efficiency scores on the entire sample of 1,048 children with and without ASD showed good model fit, with CFI = 0.947, SRMR = 0.036, and RMSEA = 0.047. The fit indices for all factor models are summarized in Table S3. All factor loadings exceeded 0.5, except for ADT (0.43). Correlations between the four cognitive domains ranged from 0.43 to 0.81 (Figure 1). A moderate correlation (0.68) between complex cognition and executive function was expected due to overlapping constructs, while social cognition and episodic memory also showed a strong correlation (0.81). The correlated 3-factor models for speed and accuracy also had acceptable fit indices (Table S3), with all the tasks loading significantly onto their corresponding factor (Supplementary Figures S1 and S2).

Confirmatory factor analysis 2: bi-factor models

Based on the established cognitive theory and the original data of the CNB (Moore et al., 2015) which indicate the presence of a general (g) factor, bi-factor models (Reise, 2012) for efficiency, accuracy, and speed scores were specified, with the cognitive

https://doi.org/10.1017/S1355617725101653 Published online by Cambridge University Press

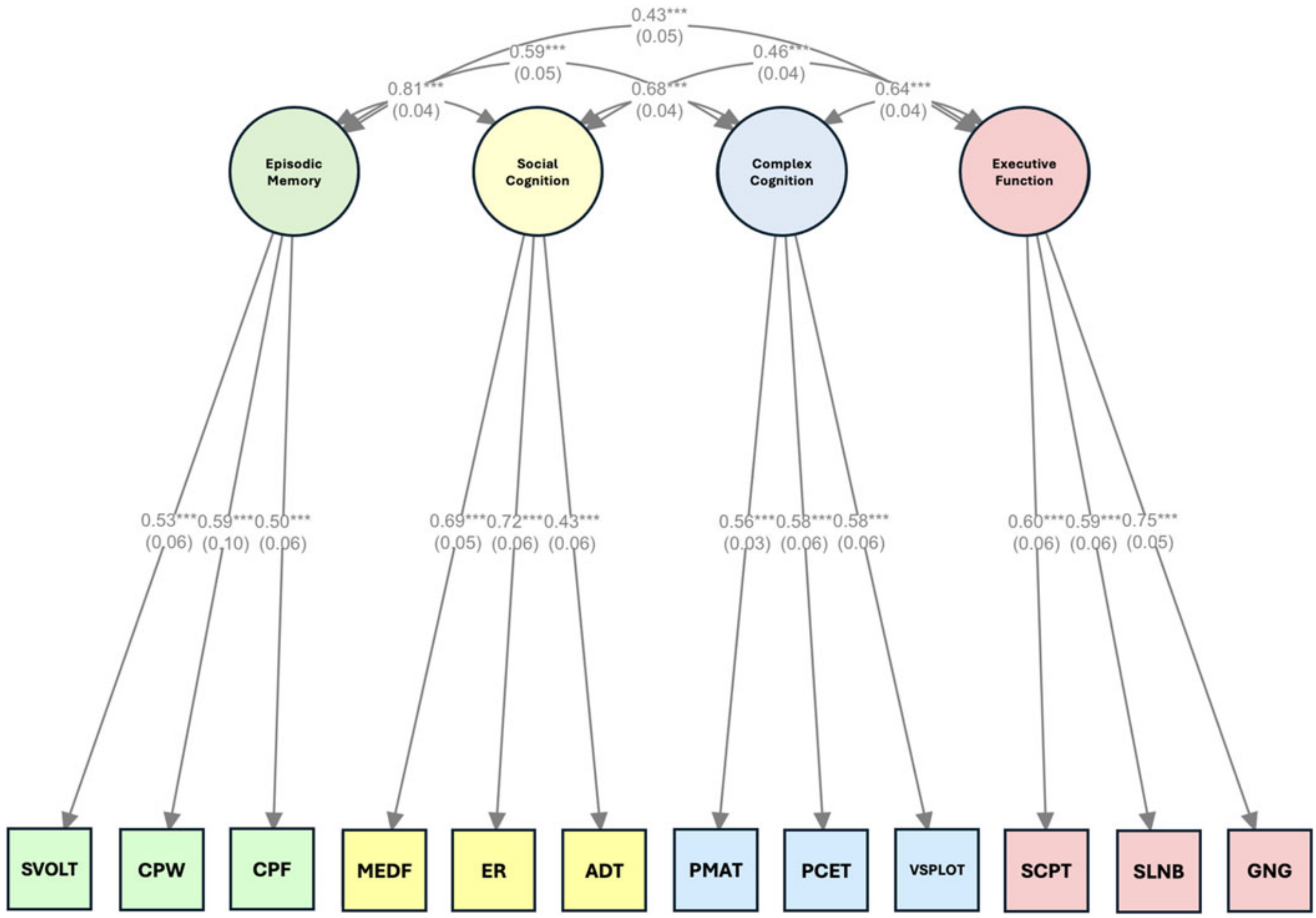

Figure 1. Correlated four-factor model for the efficiency scores of the CNB-HK.

domains and g factor being orthogonal to each other. The model of the efficiency scores fit acceptably, with a CFI = 0.947, SRMR = 0.039, and RMSEA = 0.051. While all 12 tasks loaded significantly on the g factor, tasks from episodic memory no longer loaded significantly onto their specific domain (Figure 2). A similar pattern, where certain tasks no longer loaded onto their specific domain beyond g, was observed in the bi-factor models of accuracy and speed (Supplementary Figures S3 and S4).

Exploratory factor analyses

EFAs on the entire sample were conducted and scree plot, parallel analysis, model fit indices, and theoretical interpretability were used to determine the optimal number of factors. Scree plot and parallel analysis indicated that a four-factor solution was optimal for the efficiency scores (Figure 3). We compared factor structures from a unidimensional model to a four-factor solution (Table 4). In the unidimensional model, while the loadings of all tasks ranged from 0.370 to 0.645, its fit indices were not satisfactory (Supplementary Table S3). From the 2-factor solution onward, the fit indices improved progressively, and tasks of complex cognition and executive function began to form their distinct factors. This remained true in the 4-factor solution with the fit indices CFI = 0.986, SRMR = 0.021, and RMSEA = 0.037. However, the episodic memory and social cognition tasks did not separate in the 3-factor solution, and there was a mixing of tasks loading onto each of these two factors in the 4-factor solution (Table 4). We examined whether this factor structure was specific to ASD by re-running the EFA separately for the children with and without ASD. While the 4-factor solution of the ASD group closely resembled that of the entire study population, the without ASD group had a 4-factor solution that replicated the exact theory-based factor structure of the CNB (Moore et al., 2015), with all 12 tasks loading primarily on the intended cognitive domain (Supplementary Tables S4 and S5).

The exploratory solutions for accuracy and speed are presented in Supplementary Tables S6 and S7. Three-factor solutions were optimal for both metrics, similarly aligning with the original study (Moore et al., 2015).

Factor scores generation and measurement invariance

To generate factor scores, we examined the factor determinacy (FD) (Grice, 2001), where a FD value above 0.8 is considered adequate for estimation of factor scores for each cognitive domain (Gorsuch, 1983). Across the models for efficiency, accuracy, and speed, the g factors had an FD of >0.8 in the bi-factor models, and the specific cognitive domains achieved FD values >0.8 only in the correlated-factor models (Supplementary Table S8). Hence, in the subsequent analyses, the g factor scores were generated from the bi-factor models, while the domain scores were from the correlated-factor models. Note that these

https://doi.org/10.1017/S1355617725101653 Published online by Cambridge University Press

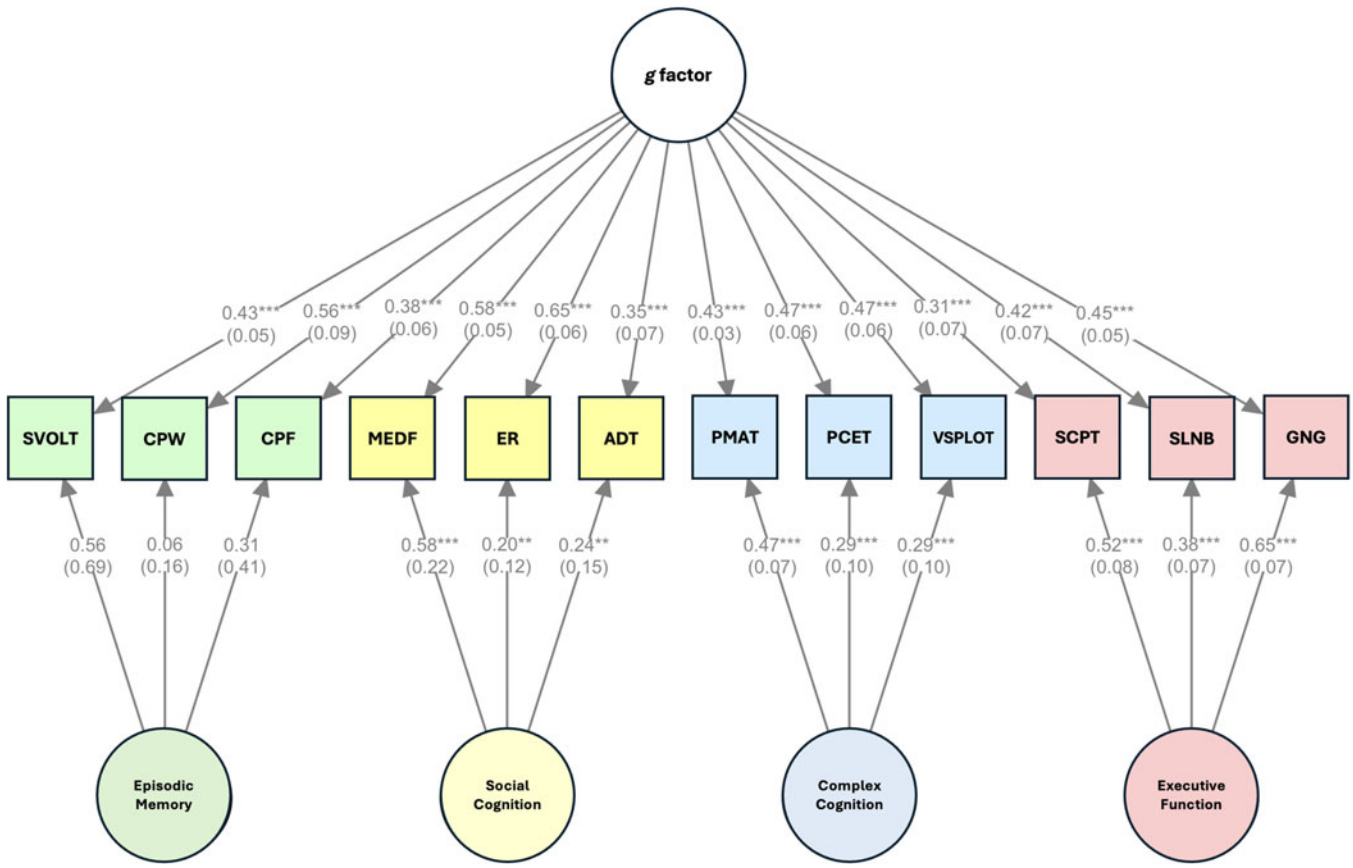

Figure 2. Bi-factor model for the efficiency scores of the CNB-HK.

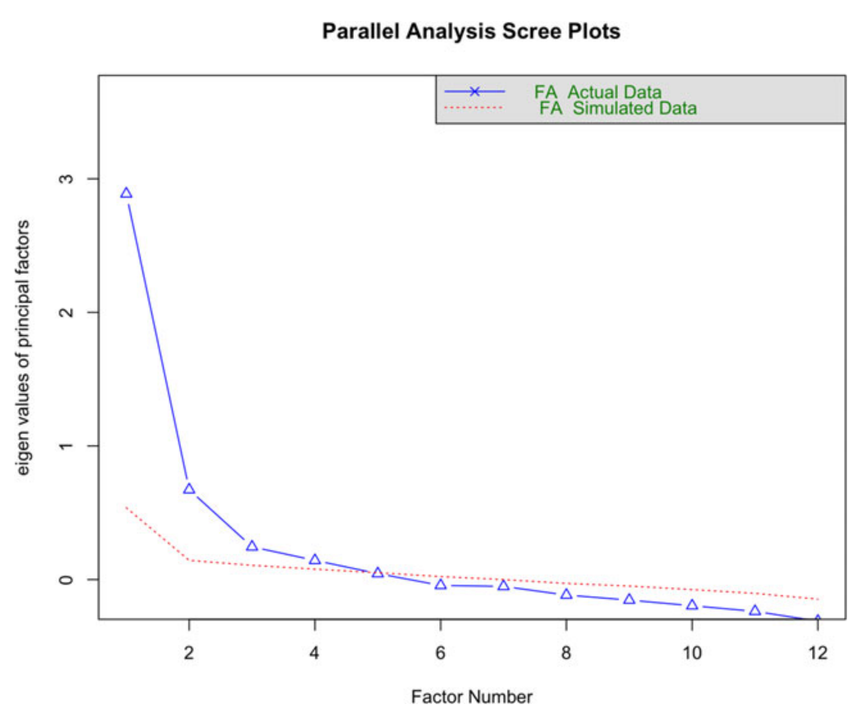

Figure 3. Scree plot and parallel analysis of the 12 CNB-HK tasks efficiency scores.

https://doi.org/10.1017/S1355617725101653 Published online by Cambridge University Press

Table 4. Exploratory factor analyses on the CNB-HK efficiency scores from unidimensional to four-factor solutions

	Unidimensional	2-Factor		3-Factor			4-Factor			
		f1	f2	f1	f2	f3	f1	f2	f3	f4
SVOLT	0.435*	0.463*		0.351*			0.582*			
CPW	0.522*	0.463*		0.541*				0.397*		
CPF	0.369*	0.514*		0.502*			0.546*			
MEDF	0.600*	0.652*		0.561*				0.445*	0.215*	
ER	0.645*	0.667*		0.722*				0.812*		
ADT	0.370*	0.507*		0.422*			0.261*			
PMAT	0.469*	0.320*	0.203*		0.565*				0.613*	
PCET	0.497*	0.330*	0.230*		0.523*				0.476*	
VSPLOT	0.503*	0.289*	0.293*		0.483*				0.449*	
SCPT	0.397*		0.604*			0.495*				0.569*
SLNB	0.477*		0.522*		0.236*	0.389*				0.429*
GNG	0.518*		0.772*			0.862*				0.826*

Factor correlations (phi matrices)

		f1	f2	f1	f2	f3	f1	f2	f3	f4
	f1	–		–			–			
	f2	0.454	–	0.575*			0.509*	–		
	f3			0.378*	0.443*	–	0.407*	0.467*	–	
	f4						0.207*	0.365*	0.469*	–

Note: Loadings < 0.20 removed. * = significant at 1% level. SVOLT = Visual Object Learning Test, CPW = Word Memory Task, CPF = Face Memory Task, MEDF = Emotional Differentiation Test, ADT = Age Differentiation Test, ER = Emotion Identification Test, PMAT = Matrix Reasoning Test, PCET = Penn Conditional Exclusion Test, VSPLOT = Line Orientation Test, SCPT = Penn Continuous Performance Test, SLNB = Letter N-Back Task, GNG = Go/No-Go Task.

scores come from different models and hence could not be used simultaneously in downstream analyses due to redundant variance in "g" and the correlated factors.

To compare the factor scores between the two populations, measurement invariance of the CNB-HK was examined. As sole use of $\Delta\chi2$ as a criterion for measurement invariance would be oversensitive given our sample size, we considered primarily the fit indices' changes at each level of configural, metric, and scalar invariance. The threshold of < -0.01 change in CFI, and <0.015 change in SRMR and RMSEA was adopted (Putnick & Bornstein, 2016). Configural and metric invariance for both the bi-factor and correlated four-factor models were established. While the $\Delta\chi2$ was significant for scalar invariance of the bi-factor model, fit indices changes (ΔCFI = −0.006; ΔSRMR = 0.002; ΔRMSEA = 0.001; $\Delta\chi2 = 19.16$, Δdf = 7, $p = 0.003$) remained below thresholds, confirming scalar invariance. However, scalar invariance was violated in the correlated four-factor model with ΔCFI = −0.057; ΔSRMR = 0.012; ΔRMSEA = 0.019; $\Delta\chi2 = 125.90$, Δdf = 8, $p < 0.001$. Univariate tests revealed that ADT, CPW, PCET, and SCPT in the correlated four-factor models significantly impacted the model fit. The scalar model was adjusted with these tasks "released" for free estimation in both groups. Subsequently, partial scalar invariance was established, where ΔCFI = −0.006; ΔSRMR = 0.001; ΔRMSEA = 0.002, despite $\Delta\chi2$ being significant ($\Delta\chi2 = 15.65$, Δdf = 4, $p = 0.003$). The results of measurement invariance analysis are summarized in Supplementary Table S3.

Correlations of factor scores with IQ and between-group comparisons

The correlations between factor scores of the CNB-HK and estimated FSIQ of the ASD population are summarized in Table 5. Generally, all factor scores correlated positively with FSIQ, with moderate strength of correlations in efficiency and accuracy factor scores and very weak correlations with speed factor scores. The Complex Cognition/Executive Function accuracy factor score had the strongest correlation with FSIQ ($\rho = 0.52$, df = 612, $p < 0.001$).

Comparison of the efficiency factor scores between children with and without ASD was conducted with a MANCOVA for the four cognitive domains and an ANCOVA for g. Both analyses were adjusted for age and sex, with the factor scores from the CFA reflected and log-transformed before analysis. Overall, the ASD group had worse cognitive performance in the four cognitive domains (Pillai's Trace = 0.08, F (5, 994) = 18.04, $p < 0.001$, $\eta p^2 = 0.08$), as well as g ($F = 19.06$, $p < 0.001$, $\eta p^2 = 0.019$). Univariate tests showed that although the ASD group had worse cognitive performance in the four specific domains, the effect sizes were different, with $\eta p^2 = 0.017$ and 0.020 respectively for episodic memory and social cognition, while complex cognition and executive function had $\eta p^2 < 0.01$ (Table 6, and Figure 4). As only partial scalar measurement invariance was established for the correlated-factor model, a reduced fully invariant correlated-factor model, with ADT, CPW, PCET, and SCPT removed (i.e. the tasks "released" for partial scalar measurement invariance), was built (Supplementary Table S9). The MANCOVA analysis was re-run and showed comparable results (Supplementary Table S10).

Discussion

In the present study, we demonstrated the feasibility of administering the CNB-HK to a cohort of children with and without ASD aged 6–12 years. Factor structure of the CNB-HK replicated that of the original CNB, and task performance could be summarized into factor scores that correlated positively with FSIQ. Thus, the CNB-HK is a valid instrument for assessing multi-domain cognitive function in children with ASD. Compared to children without ASD, children with ASD performed more poorly across all cognitive domains.

Feasibility of the CNB-HK as a multi-domain cognitive assessment for children with ASD

The merits of the CNB-HK include its automated scoring, ease of administration, and ability to assess major neurocognitive domains efficiently in one hour. Additionally, the CNB-HK assesses

https://doi.org/10.1017/S1355617725101653 Published online by Cambridge University Press

Table 5. Spearman's rho correlation matrix for age, FSIQ, and CNB-HK factor scores (Efficiency, accuracy, speed) in the ASD group

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Age		—														
2. FSIQ		−0.096	—													
Efficiency	3. Episodic memory	0.066	0.425***	—												
	4. Social cognition	0.099*	0.45***	0.937***	—											
	5. Complex cognition	0.05	0.499***	0.77***	0.814***	—										
	6. Executive function	−0.01	0.399***	0.583***	0.593***	0.786***	—									
	7. g	0.059	0.492***	0.946***	0.957***	0.909***	0.73***	—								
Accuracy	8. Episodic memory	0.146***	0.427***	0.651***	0.628***	0.522***	0.443***	0.65***	—							
	9. Social cognition	0.197***	0.405***	0.582***	0.672***	0.509***	0.418***	0.625***	0.88***	—						
	10. Complex cognition & executive function	0.088*	0.521***	0.468***	0.524***	0.667***	0.616***	0.603***	0.729***	0.748***	—					
	11. *g*	0.107**	0.462***	0.619***	0.65***	0.569***	0.533***	0.667***	0.947***	0.922***	0.828***	—				
Speed	12. Episodic memory	−0.061	0.112*	0.597***	0.501***	0.446***	0.29***	0.512***	−0.124**	−0.143***	−0.101*	−0.135***	—			
	13. Deliberative	−0.055	0.188***	0.636***	0.613***	0.583***	0.402***	0.622***	−0.039	−0.09*	−0.046	−0.052	0.896***	—		
	14. Vigilance	−0.019	0.15**	0.32***	0.303***	0.452***	0.647***	0.4***	0.015	−0.004	0.067	−0.01	0.404***	0.472***	—	
	15. g	−0.051	0.225***	0.682***	0.656***	0.617***	0.443***	0.671***	0.018	−0.034	0.016	0.009	0.893***	0.989***	0.476***	—

Note. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, FSIQ = Full-Scale Intelligence Quotient, ASD = Autism Spectrum Disorder.

cognitive domains critical to ASD: social cognition, a domain linked to the social communication deficit; processing speed, which is often impaired in ASD (Zapparrata et al., 2023); and executive function, which is associated with prognosis (Kenny et al., 2019). These features support the CNB-HK's ecological validity for ASD research.

Most children with ASD completed the battery, with 72.7% providing valid data for all 12 tasks and 91.1% for at least 11 tasks, indicating broad applicability for school-age ASD. However, a notable limitation was the higher invalid data rate in 6-year-old ASD, with only 53.2% completing all 12 tasks in this age group. Despite this, 70.3% and 80.2% of 6-year-olds could provide respectively valid data for at least 11 and 10 tasks. With three tasks per cognitive domain, the CNB-HK allows missing data from a single task to be managed using imputations. While this supports applicability for 6-year-olds, imputation may introduce assumptions affecting result precision. The invalid data among 6-year-olds was particularly notable for executive function tasks (SLNB, GNG, CPT), aligning with findings from the NIH-TCB, where only 57% of participants with ASD, including preschoolers (as young as 3 years) and individuals with intellectual disability (ID), completed the four tasks that assessed executive function (Jones et al., 2022). These findings caution the applicability of existing executive function tasks, including those in the CNB-HK, to populations with younger mental ages. Adaptations such as using a gamified platform, shortening, and adjusting the difficulty of the executive tasks may be necessary for preschoolers to children around 6 years old (Moore et al., 2023). In summary, we recommend a minimum age of 6 years for CNB-HK administration, with the caveat that executive function tasks may require tailoring for younger children.

Factor structure of the CNB-HK and factor scores generation

With the necessary task-level changes for adaptation of the CNB to the Chinese population, the 4-factor CFA models (bi-factor and correlated-factor models) fit well with the data when all tasks were assigned to their respective designed factors, consistent with other CNB adaptations for youth with HIV in Botswana (Pelt et al., 2022), Xhosa-speaking adults with schizophrenia in South Africa (Scott et al., 2021), and healthy individuals in Portugal (da Motta et al., 2021). This demonstrates that the factor structure of the CNB is stable cross-culturally for trans-diagnostic comparisons. Curiously, in the EFA, tasks within episodic memory and social cognition did not separate in the ASD group, and the two domains showed a high correlation. At first glance, a 3-factor solution may appear sufficient in the population with ASD. Yet, the observation can be explained by the core deficits of ASD leading to similar task performances in the two domains, while considering task impurity and shared task demand in cognitive assessment. For instance, word memory (CPW) captures language impairment which is causally linked to social deficits (Constantino et al., 2004), while face memory (CPF) and social cognition tasks (ADT, ER, MEDF) all require facial stimuli processing, a key component of social cognition (Haxby et al., 2002). Additionally, this correlation may be related to shared neural mechanisms in the temporo-limbic system, with evidence suggesting that social interaction and memory are supported by the hippocampus in ASD (Banker et al., 2021; Roesler et al., 2021). These biological mechanisms warrant further investigation to clarify their role in the observed correlation. The EFA results did not invalidate the ability of the tasks to capture the performance of their intended domains,

Table 6. MANCOVA on cognitive domain factor efficiency scores between children with and without ASD

				Multivariate tests			
		value	*F*	*df1*	*df2*	*p*	ηp^2
Group	Pillai's trace	0.08	18.04	5	994	<0.001***	0.08
Univariate tests							
	Dependent Variable	SS	*df*	MS	*F*	*p*	ηp^2
Group	Episodic memory	0.39	1	0.39	27.79	<.001***	0.017
	Social cognition	0.51	1	0.51	33.58	<.001***	0.02
	Complex cognition	0.09	1	0.09	6.81	0.009**	6.00E-03
	Executive function	0.11	1	0.11	7.48	0.006**	9.00E-03
Residuals	Episodic memory	13.98	998	0.01			
	Social cognition	15.18	998	0.02			
	Complex cognition	13.54	998	0.01			
	Executive function	15.13	998	0.02			

Note. Results were controlled for age and sex. $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. ASD = Autism Spectrum Disorder.

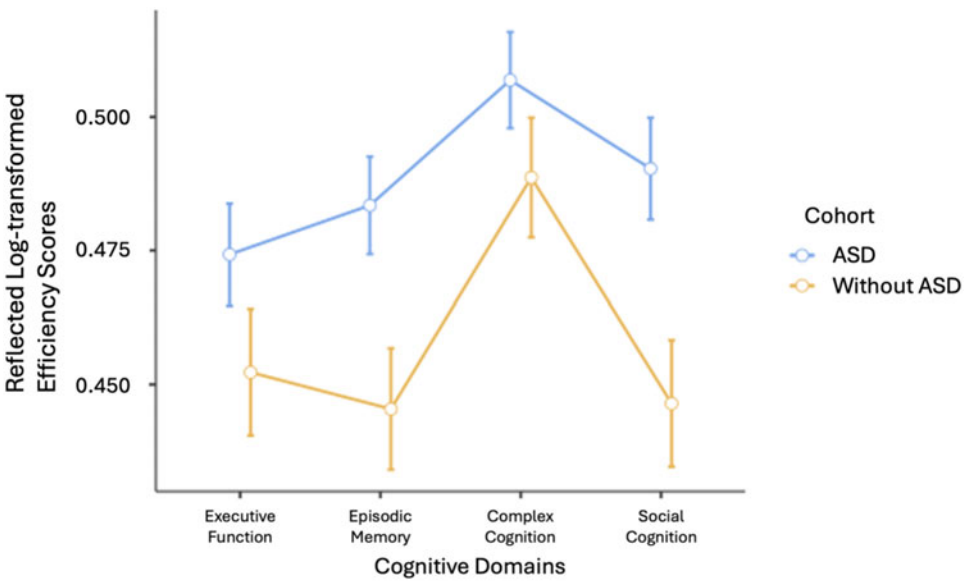

Figure 4. Comparison of the domain-specific factor efficiency scores of the CNB-HK between children with and without ASD.
Note: Error bars represent 95% confidence intervals.

given that in the correlated-factor CFA model, all the tasks still loaded significantly onto their intended domain.

When generating factor scores from factor analysis, the loadings and number of the indicators (i.e. tasks) for a factor are critical to whether the score is well-determined. In the bi-factor CFA model, when all 12 tasks contributed as indicators of the g factor, little variance was left for the specific cognitive domains. Consequently, while the g factor had adequate factor determinacy, the domain-specific factors in the bi-factor model did not; they were sufficiently determined only in the correlated-factor model. Although the orthogonality of the factors within the bi-factor model can be advantageous for examining specific cognitive effect, achieving determinacy for domain-specific factors in the bi-factor model would require substantially increasing the number of tasks per domain, which would prolong and compromise the feasibility of the battery. Conversely, even though the domain-specific factor scores in the correlated-factor CFA model were well-determined, their specificity was compromised due to shared variance among the scores. The phenomenon of low factor loadings for specific domains after accounting for the g factor, as well as the trade-off between specificity and determinacy were also observed in the original study of the CNB (Moore et al., 2015). In optimal circumstances, a bi-factor model provides well-determined general and specific factor scores, allowing one to study all factors simultaneously – e.g., the general and specific factor scores can be included in a regression model together – without the cross-contamination between the specific factors and the general factor (as in the correlated-traits model). Here, however, specific factors of the bi-factor model had such weak loadings that they were too poorly determined to generate reliable scores. We therefore had to use both models, each for a different purpose: the correlated-factor model was our only option for examining domain-specific cognitive abilities, such as targeting specific deficits in ASD (e.g., episodic memory or social cognition), due to its adequate factor determinacy. In contrast, the bi-factor model does provide better estimates of general cognitive performance (g factor) than, for example, simply taking the average of the correlated-factor scores; we therefore used the bi-factor model for overall “g” scores.

https://doi.org/10.1017/S1355617725101653 Published online by Cambridge University Press

Correlation of the CNB-HK factor scores with FSIQ

The generalized positive correlation between all factor scores (g and domain-specific accuracy, speed, and efficiency scores) and estimated FSIQ provided construct validity for the CNB-HK as a cognitive assessment tool for children with ASD. As both the WISC-IV (HK) Short Form IQ subtests include complex cognition and executive tasks, FSIQ exhibited the strongest correlation with the accuracy factor score of these two domains. Except for the digit symbol substitution task, the subtests of the WISC-IV did not consider the speed of response for scoring, which may explain the weak correlations between FSIQ and the speed factor scores.

Measurement invariance and comparison between children with and without ASD

Measurement invariance is an important aspect to ensuring comparability of the factor scores between the groups. Briefly, configural invariance confirms that the same factor structure (e.g., cognitive domains) holds across groups, while metric invariance ensures that factor loadings are equivalent. Scalar invariance verifies that item intercepts are equivalent to enable valid comparisons of group means (Putnick & Bornstein, 2016). In our study, even though configural and metric invariance were both established, we could only achieve partial scalar invariance of the correlated-factor CFA model. Although partial scalar invariance may introduce estimation bias in between-group comparisons, the extent to which such violations preclude valid comparisons remains under investigation (Davidov et al., 2014; Putnick & Bornstein, 2016). To examine spurious results from such violations, a sensitivity analysis was done with a reduced and fully invariant model. As similar results were observed, the impact of the partial scalar invariance may not be significant.

Although the children with ASD in this study were recruited from a specialist clinic, they represented a spectrum of clinical severity (with and without ADHD), and varying intelligence levels (a mean FSIQ of 97.5 and SD of 16.10). The MANCOVA showed that, at the factor level, the ASD group had a generalized cognitive impairment across all cognitive domains and the g factor. However, the effect sizes of the cognitive impairments vary. Social cognition, a core deficit of ASD, exhibited the largest effect size, followed by memory, while effect sizes for complex cognition and executive function were smaller, indicating less pronounced impairments in these areas. This pattern of varying effect sizes was consistent with the findings from the ASD group in the PNC (Wong et al., 2023). Although executive dysfunction is common in individuals with ASD (Sadozai et al., 2024), the small effect size for executive function differences between children with and without ASD ($\eta p2 = 9.00E\text{-}03$) suggests some overlap in performance distributions. This indicates that some children with ASD likely possess intact executive function, despite group-level impairments. Well-developed executive function has been conceptualized as protective in neurodevelopmental conditions owing to its postulated compensatory mechanism against other deficits (Johnson, 2012). It is also a major determining factor for co-occurring psychopathology and functional level in youth with ASD (Wong et al., 2023), and their longitudinal outcomes (Kenny et al., 2019). Given the variability of executive function and clinical profile within our cohort, further exploration of their interplay may shed light on the linkage of cognitive function and clinical heterogeneity of ASD. Previous studies have also identified clinical correlates of social cognition, memory, and complex cognition in ASD (Bishop-Fitzpatrick et al., 2017; Edirisooriya et al., 2021; Toichi, 2008). To this end, the CNB-HK allows concurrent examination of multiple cognitive domains to pinpoint specific correlates and predictors of clinical outcomes.

Limitations

There are several limitations in the present study. First, since the ID and younger populations were not included, the results from this study may not be generalizable to these populations. As discussed, the CNB-HK is feasible for individuals with a mental age of six years or above. Further adaptation of the battery for the younger and ID population is needed. Second, the factor determinacy was inadequate for domain-specific factor scores from the bi-factor model, such that only the scores from the correlated-factor model could be used. This limits the specificity of the analysis using the scores with shared variance. Third, the CNB-HK was not fully invariant at the scalar level, and the effect of this non-invariance on the subsequent between-group comparisons is not fully understood. Fourth, although the CNB-HK factor scores exhibited a generalized positive correlation with estimated IQ, they have not been validated against established neuropsychological measures of the respective cognitive domains. Finally, females were relatively underrepresented among children with ASD, and the intrinsic sex differences in cognitive ability should be addressed for in future studies.

Conclusion

Our study showed that the CNB-HK is a feasible and valid multi-domain cognitive assessment battery for normal IQ school-age individuals with and without ASD. The stable and replicable factor structure enables cross-cultural and trans-diagnostic comparison of cognitive functions in ASD with other conditions. Future research should examine clinical correlates of cognitive profiles to guide understanding of the heterogeneity of ASD.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S1355617725101653.

Acknowledgments. None.

Funding statement. This work was support by The D.H. Chen Foundation.

Competing interests. The authors declare no conflict of interest.

References

Banker, S. M., Gu, X., Schiller, D., & Foss-Feig, J. H. (2021). Hippocampal contributions to social and cognitive deficits in autism spectrum disorder. *Trends in Neurosciences*, *44*, 793–807.

Bassett, D. S., & Sporns, O. (2017). Network neuroscience. *Nature Neuroscience*, *20*, 353–364. https://doi.org/10.1038/nn.4502

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*, 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Bhatia, T., Mazumdar, S., Wood, J., He, F., Gur, R. E., Gur, R. C., Nimgaonkar, V. L., & Deshpande, S. N. (2017). A randomised controlled trial of adjunctive yoga and adjunctive physical exercise training for cognitive dysfunction in schizophrenia. *Acta Neuropsychiatrica*, *29*, 102–114. https://doi.org/10.1017/NEU.2016.42

Bishop-Fitzpatrick, L., Mazefsky, C. A., Eack, S. M., & Minshew, N. J. (2017). Correlates of social functioning in autism spectrum disorder: The role of social cognition. *Research in Autism Spectrum Disorders*, *35*, 25–34. https://doi.org/10.1016/j.rasd.2016.11.013

Calkins, M. E., Merikangas, K. R., Moore, T. M., Burstein, M., Behr, M. A., Satterthwaite, T. D., Ruparel, K., Wolf, D. H., Roalf, D. R., Mentch, F. D., Qiu, H., Chiavacci, R., Connolly, J. J., Sleiman, P. M. A., Gur, R. C., Hakonarson, H., & Gur, R. E. (2015). The Philadelphia Neurodevelopmental Cohort: Constructing a deep phenotyping collaborative. *Journal of Child Psychology and Psychiatry*, *56*, 1356–1369. https://doi.org/10.1111/jcpp.12416

Chan, S. S. M., Wong, O. W. H., Hussain, S., Tsoi, K. K. F., Ma, K. K. Y., Chau, S. W. H., Ma, S. L., Lai, K. Y. C., Chu, W. C. W., Lo, H. H. L., Ho, S. W. S., Leung, C. C., Yiu, K. K. L., So, S. H. W., Sham, P. C., Hung, S. F., & Leung, P. W. L. (2025). Twelve-month prevalence of DSM-5 mental disorders and the psychosocial correlates– A child and adolescent psychiatric epidemiologic survey in Hong Kong SAR. *The Lancet Regional Health – Western Pacific*, *57*, 101533.https://doi.org/10.1016/j.lanwpc.2025.101533

Constantino, J. N., Gruber, C. P., Davis, S., Hayes, S., Passanante, N., & Przybeck, T. (2004). The factor structure of autistic traits. *Journal of Child Psychology and Psychiatry*, *45*, 719–726. https://doi.org/10.1111/J.1469-7610.2004.00266.X

da Motta, C., Carvalho, C. B., Castilho, P., & Pato, M. T. (2021). Assessment of neurocognitive function and social cognition with computerized batteries: Psychometric properties of the Portuguese PennCNB in healthy controls. *Current Psychology*, *40*, 4851–4862. https://doi.org/10.1007/s12144-019-00419-2

Davidov, E., Meuleman, B., Cieciuch, J., Schmidt, P., & Billiet, J. (2014). Measurement equivalence in cross-national research. *Annual Review of Sociology*, *40*, 55–75. https://doi.org/10.1146/annurev-soc-071913-043137

Edirisooriya, M., Dykiert, D., & Auyeung, B. (2021). IQ and internalising symptoms in adolescents with ASD. *Journal of Autism and Developmental Disorders*, *51*, 3887–3904. https://doi.org/10.1007/s10803-020-04810-y

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates. https://doi.org/10.4324/9780203781098

Grice, J. W. (2001). Computing and evaluating factor scores. *Psychological Methods*, *6*, 430–450. https://doi.org/10.1037/1082-989X.6.4.430

Grove, J., Ripke, S., Als, T. D., Mattheisen, M., Walters, R. K., Won, H., Pallesen, J., Agerbo, E., Andreassen, O. A., Anney, R., Awashti, S., Belliveau, R., Bettella, F., Buxbaum, J. D., Bybjerg-Grauholm, J., Bækvad-Hansen, M., Cerrato, F., Chambert, K., Christensen, J. H., … Børglum, A. D. (2019). Identification of common genetic risk variants for autism spectrum disorder. *Nature Genetics*, *51*, 431–444. https://doi.org/10.1038/s41588-019-0344-8

Gur, R. C., Ragland, J. D., Moberg, P. J., Turner, T. H., Bilker, W. B., Kohler, C., Siegel, S. J., & Gur, R. E. (2001). Computerized neurocognitive scanning: I methodology and validation in healthy people. *Neuropsychopharmacology*, *25*, 766–776. https://doi.org/10.1016/S0893-133X(01)00278-0

Gur, R. C., Richard, J., Hughett, P., Calkins, M. E., Macy, L., Bilker, W. B., Brensinger, C., & Gur, R. E. (2010). A cognitive neuroscience-based computerized battery for efficient measurement of individual differences: Standardization and initial construct validation. *Journal of Neuroscience Methods*, *187*, 254–262. https://doi.org/10.1016/j.jneumeth.2009.11.017

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2002). Human neural systems for face recognition and social communication. *Biological Psychiatry*, *51*, 59–67. https://doi.org/10.1016/S0006-3223(01)01330-0

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55. https://doi.org/10.1080/10705519909540118

Izgi, B., Moore, T. M., Yalcinay-Inan, M., Port, A. M., Kuscu, K., Gur, R. C., & Eser, H. Y. (2022). Test–retest reliability of the Turkish translation of the Penn computerized neurocognitive battery. *Applied Neuropsychology: Adult*, *29*, 1258–1267. https://doi.org/10.1080/23279095.2020.1866572

Johnson, M. H. (2012). Executive function and developmental disorders: The flip side of the coin. *Trends in Cognitive Sciences*, *16*, 454–457. https://doi.org/10.1016/j.tics.2012.07.001

Jones, D. R., Dallman, A., Harrop, C., Whitten, A., Pritchett, J., Lecavalier, L., Bodfish, J. W., & Boyd, B. A. (2022). Evaluating the feasibility of the NIH toolbox cognition battery for autistic children and adolescents. *Journal of Autism and Developmental Disorders*, *52*, 689–699. https://doi.org/10.1007/S10803-021-04965-2

Joseph, R. M., Tager-Flusberg, H., & Lord, C. (2002). Cognitive profiles and social-communicative functioning in children with autism spectrum disorder. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *43*, 807–821. https://doi.org/10.1111/1469-7610.00092

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.

Kenny, L., Cribb, S. J., & Pellicano, E. (2019). Childhood executive function predicts later autistic features and adaptive behavior in young autistic people: A 12-year prospective study. *Journal of Abnormal Child Psychology*, *47*, 1089–1099. https://doi.org/10.1007/s10802-018-0493-8

Leung, C. N. W., Leung, C. S. Y., Chan, R. W. S., & Leung, P. W. L. (2023). Can the UK AQ-10 be applicable to Chinese samples with autism spectrum disorder in Hong Kong? Cross-cultural similarities and differences. *Autism Research*, *16*, 302–314. https://doi.org/10.1002/AUR.2847

Lombardo, M. V., Lai, M. C., & Baron-Cohen, S. (2019). Big data approaches to decomposing heterogeneity across the autism spectrum. *Molecular Psychiatry*, *24*, 1435–1450. https://doi.org/10.1038/s41380-018-0321-0

Menon, V., & D'Esposito, M. (2021). The role of PFC networks in cognitive control and executive function. *Neuropsychopharmacology*, *47*, 90–103. https://doi.org/10.1038/s41386-021-01152-w

Moore, T. M., Di Sandro, A., Scott, J. C., Lopez, K. C., Ruparel, K., Njokweni, L. J., Santra, S., Conway, D. S., Port, A. M., D'Errico, L., Rush, S., Wolf, D. H., Calkins, M. E., Gur, R. E., & Gur, R. C. (2023). Construction of a computerized adaptive test (CAT-CCNB) for efficient neurocognitive and clinical psychopathology assessment. *Journal of Neuroscience Methods*, *386*, 109795. https://doi.org/10.1016/J.JNEUMETH.2023.109795

Moore, T. M., Reise, S. P., Gur, R. E., Hakonarson, H., & Gur, R. C. (2015). Psychometric properties of the Penn computerized neurocognitive battery. *Neuropsychology*, *29*, 235–246. https://doi.org/10.1037/neu0000093

Pelt, A. E. V., Scott, J. C., Morales, K. H., Matshaba, M., Gur, R. C., Tshume, O., Thuto, B., Lowenthal, E. D., & Moore, T. M. (2022). Structural validity of a computerized neurocognitive battery for youth affected by human immunodeficiency virus in Botswana. *Psychological Assessment*, *34*, 139–146. https://doi.org/10.1037/PAS0001066

Van Pelt, A. E., Moore, T. M., Scott, J. C., Phoi, O., Mbakile-Mahlanza, L., Morales, K. H., Gur, R. C., Rampa, S., Matshaba, M., & Lowenthal, E. D. (2022). Predictive validity of a computerized battery for identifying neurocognitive impairments among children living with HIV in Botswana. *AIDS and Behavior*, *26*, 2758–2767. https://doi.org/10.1007/s10461-022-03620-w

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

Reise, S. P. (2012). The rediscovery of bi-factor measurement models. *Multivariate Behavioral Research*, *47*, 667–696. https://doi.org/10.1080/00273171.2012.715555

Revelle, W. (2020). *Psych: Procedures for psychological, psychometric, and personality research* (R package version 2.9).

Roesler, R., Parent, M. B., LaLumiere, R. T., & McIntyre, C. K. (2021). Amygdala-hippocampal interactions in synaptic plasticity and memory formation. *Neurobiology of Learning and Memory*, *184*, 107490. https://doi.org/10.1016/J.NLM.2021.107490

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*, 1–36. https://doi.org/10.18637/JSS.V048.I02

Sadozai, A. K., Sun, C., Demetriou, E. A., Lampit, A., Munro, M., Perry, N., Boulton, K. A., & Guastella, A. J. (2024). Executive function in children with neurodevelopmental conditions: A systematic review and meta-analysis. *Nature Human Behaviour*, *2024*, 1–10. https://doi.org/10.1038/s41562-024-02000-9

Scott, J. C., Moore, T. M., Stein, D. J., Pretorius, A., Zingela, Z., Nagdee, M., Ngqengelele, L., Campbell, M., Sibeko, G., King, M. C., McClellan, J. M., Port, A. M., Jackson, C., Ruparel, K., Susser, E. S., & Gur, R. C. (2021). Adaptation and validation of a computerized neurocognitive battery in the Xhosa of South Africa. *Neuropsychology*, *35*, 581. https://doi.org/10.1037/NEU0000742

Service, S. K., Vargas Upegui, C., Castaño Ramírez, M., Port, A. M., Moore, T. M., Munoz Umanes, M., Agudelo Arango, L. G., Díaz-Zuluaga, A. M., Melo Espejo, J., López, M. C., Palacio, J. D., Ruiz Sánchez, S., Valencia, J., Teshiba, T. M., Espinoza, A., Olde Loohuis, L., De la Hoz Gomez, J., Brodey, B. B.,

https://doi.org/10.1017/S1355617725101653 Published online by Cambridge University Press

Sabatti, C., . . . Freimer, N. B. (2020). Distinct and shared contributions of diagnosis and symptom domains to cognitive performance in severe mental illness in the Paisa population: A case–control study. *The Lancet Psychiatry*, *7*, 411–419. https://doi.org/10.1016/S2215-0366(20)30098-5

Shaffer, D., Fisher, P., Lucas, C. P., Dulcan, M. K., & Schwab-Stone, M. E. (2000). NIMH Diagnostic Interview Schedule for Children Version IV (NIMH DISC-IV): Description, differences from previous versions, and reliability of some common diagnoses. *Journal of the American Academy of Child and Adolescent Psychiatry*, *39*, 28–38. https://doi.org/10.1097/00004583-200001000-00014

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*, *25*, 173–180. https://doi.org/10.1207/s15327906mbr2502_4

Su, Q., Wong, O. W. H., Lu, W., Wan, Y., Zhang, L., Xu, W., Li, M. K. T., Liu, C., Cheung, C. P., Ching, J. Y. L., Cheong, P. K., Leung, T. F., Chan, S., Leung, P., Chan, F. K. L., & Ng, C. (2024). Multikingdom and functional gut microbiota markers for autism spectrum disorder. *Nature Microbiology*, *9*, 1–12. https://doi.org/10.1038/s41564-024-01739-1

Toichi, M. (2008). Episodic memory, semantic memory and self-awareness in high-functioning autism. In J. Boucher & D. Bowler (Eds.), *Memory in autism* (pp. 143–165). Cambridge University Press. https://doi.org/10.1017/CBO9780511490101.010

Van Widenfelt, B. M., Treffers, P. D. A., De Beurs, E., Siebelink, B. M., & Koudijs, E. (2005). Translation and cross-cultural adaptation of assessment instruments used in psychological research with children and families. *Clinical Child and Family Psychology Review*, *8*, 135–147. https://doi.org/10.1007/S10567-005-4752-1

Wan, Y., Wong, O. W. H., Tun, H. M., Su, Q., Xu, Z., Tang, W., Ma, S. L., Chan, S., Chan, F. K. L., & Ng, S. C. (2024). Fecal microbial marker panel for aiding diagnosis of autism spectrum disorders. *Gut Microbes*, *16*, 2418984. https://doi.org/10.1080/19490976.2024.2418984

Wechsler, D. (2010). *Wechsler Intelligence Scale for Children*. 4th ed, short-form. Hong Kong, China: King-May Psychological Assessment.

Wong, O. W., Chan, S. S., Chau, S. W., Chu, W. C., Ho, C. S., Ho, S. W., Hung, S. F., Hussain, S., Lai, K. Y., Lam, A. M., Lo, H. H., Ma, K. K., Ma, S. L., Mo, F. Y., Sham, P. C., Shea, C. K., So, S. H., Tsoi, K. K., & Leung, P. W. (2025). Autism epidemiology in Hong Kong children and youths aged 6–17: Implications on autism screening and sex differences in the community. *Autism*, *29*, 2872–2884. https://doi.org/10.1177/13623613251360269

Wong, O. W. H., Barzilay, R., Lam, A. M. W., Chan, S., Calkins, M. E., Gur, R. E., & Gur, R. C. (2023). Executive function as a generalized determinant of psychopathology and functional outcome in school-aged autism spectrum disorder: A case-control study. *Psychological Medicine*, *53*, 4788–4798. https://doi.org/10.1017/S0033291722001787

Zapparrata, N. M., Brooks, P. J., & Ober, T. M. (2023). Slower processing speed in autism spectrum disorder: A meta-analytic investigation of time-based tasks. *Journal of Autism and Developmental Disorders*, *53*, 4618–4640. https://doi.org/10.1007/S10803-022-05736-3