BMC Psychology

RESEARCH

Open Access

The development and validation of a tablet-based assessment battery of general cognitive ability

Dong-ni Pan[1], Hailun Xie[2], Yanjia Zeng[1], Yixiang Zhou[1], Cuizhu Lin[1], Xin Ma[1], Juejing Ren[2], Yuanyun Jiao[2], Yingying Wu[2], Wei Wei[3] and Gui Xue[2,4*]

Abstract

Background Traditional cognitive assessments, often reliant on paper-and-pencil tests and professional evaluators, suffer from subjectivity and limited result discrimination. This study introduces the Baguan Online Cognitive Assessment System (BOCAS), a tablet-based system that evaluates both general cognitive ability (GCA) and domain-specific functions across six domains: sensory-motor skills, processing speed, sustained attention, working memory, cognitive flexibility, and spatial ability.

Methods BOCAS was validated with 151 healthy Chinese adults aged 18–40. Reliability was assessed through internal consistency and test-retest reliability. Factor analysis and confirmatory factor analysis (CFA) were used to validate the model. The GCA score was correlated with the Raven IQ test and self-reported cognitive flexibility, and its relationship with negative emotions (depression and anxiety) was examined.

Results BOCAS showed satisfactory reliability, with internal consistency ranging from 0.712 to 0.846 and test-retest reliability from 0.56 to 0.71. Factor analysis revealed a common factor explaining 40% of the variance, and CFA indicated a good model fit (χ^2/df = 1.81; CFI = 0.932). The GCA score strongly correlated with the Raven IQ test (r = 0.58) and was related to self-reported cognitive flexibility and negative emotions.

Conclusion BOCAS offers a digital solution for cognitive assessment, providing automated, remote, and precise evaluations. It demonstrates reliability, validity, and potential for use in clinical and research settings.

Keywords Automated electronic test, Chinese digital cognitive test, Cognitive domain, General cognitive assessment, Neuropsychological tests

*Correspondence:
Gui Xue
gxue@bnu.edu.cn
[1]School of Psychology, Beijing Language and Culture University, Beijing 100083, China
[2]Beijing Infinite Brain Technology Co., Ltd, Beijing 100022, PR China
[3]Affiliated Mental Health Center & Hangzhou Seventh People's Hospital, School of Medicine, Zhejiang University, Hangzhou 310058, PR China
[4]State Key Laboratory Cognitive Neuroscience and Learning, Beijing Normal University, Beijing 100875, China

Introduction

Cognitive ability, which encompasses an individual's aptitude in activities such as perception, thinking, memory, learning, and problem-solving, plays a pivotal role in their educational, professional, and social endeavors [1]. Introduced by Spearman in the early 20th century, the concept of general cognitive ability, often termed the "g factor," suggests that a core, universal cognitive ability affects performance across an array of cognitive tasks [2]. This idea has been validated in multiple studies, illustrating that a single factor—general cognitive ability—consistently accounts for variability in test scores across varied cognitive tasks [3]. The concept holds significant implications, especially in education and human resource management, where identifying an individual's cognitive skills is closely tied to understanding their learning potential [4, 5]. Moreover, assessing cognitive function is crucial for gauging academic and professional performance [6].

In addition to educational and occupational contexts, the assessment of cognitive ability is crucial in clinical environments. Several conditions, including abnormal aging [7], psychological and mental disorders such as depression [8] and schizophrenia [9], as well as neurodevelopmental disorders like autism [10] and attention deficit hyperactivity disorder (ADHD) [11], can result in cognitive impairment. The screening, monitoring, and intervention of these conditions necessitate accurate tools for evaluating cognitive function [12, 13].

While there is a broad demand for cognitive assessment tools, there's a notable scarcity of comprehensive assessments suitable for both clinical and educational settings. Traditional IQ tests, though invaluable, frequently lack the adaptability essential for clinical evaluations. These tests often focus on abstract components indicative of general cognitive ability, potentially overlooking specific cognitive deficits vital for clinical diagnoses. Clinical tools, such as the Mini-Mental State Examination (MMSE) and the Montreal Cognitive Assessment (MoCA), are undoubtedly useful. However, they are tied down by specific cutoff scores that differentiate normal from abnormal performance [14]. Such tests don't offer continuous, nuanced data for individual cognitive dimensions, restricting their applicability in clinical and broader populations. Moreover, traditional cognitive assessments often suffer from issues like low score discrimination and unsuitable normative standards for score conversions [15]. Furthermore, these tests usually rely on paper-and-pencil methods and professional assessors, potentially introducing subjective bias into the results [16].

With the emergence of digital cognitive assessments, innovative solutions to longstanding challenges have been introduced. These digital tools offer precise control over stimuli, facilitate the collection of in-depth performance data, reduce labor costs, and optimize data storage. Instruments such as the CogState test [17], the Cambridge Neuropsychological Test Automated Battery (CANTAB), and the National Institutes of Health's NIH Toolbox have become pivotal in both clinical and educational spheres. Furthermore, the internet-based cognitive testing platform, Lumosity, with its remote real-time testing capabilities, has amassed a comprehensive dataset of over 35 million subjects across multiple countries [18]. These digital tools offer versatile assessments for various cognitive domains and have the potential to bridge the gap between different application scenarios.

Within the Chinese context, the deployment of these assessments encounters distinctive challenges. While cognitive assessments are not inherently culture-specific, they often require careful adaptation and validation to ensure their suitability for diverse cultural contexts, particularly when it comes to translating test content into Chinese. Moreover, many digital evaluations lack normative data tailored for the Chinese demographic, which can introduce potential assessment biases. Furthermore, many digital assessments lack normative data specific to the Chinese population, potentially leading to assessment biases. Existing electronic assessments often focus on single tasks, developed piecemeal to meet specific goals [1], rather than providing a comprehensive assessment of general cognitive ability. This study endeavors to tackle these concerns by introducing a Tablet-based General Cognitive Ability Test crafted for the Chinese milieu, termed the Baguan Online Cognitive Assessment System (BOCAS). BOCAS comprises an array of tasks spanning Sensory-Motor, Processing Speed, Attention, Cognitive Flexibility, Working Memory, and Spatial Ability—fundamental facets of cognitive prowess. Each domain is not only underpinned by theoretical relevance [3] but also finds pragmatic utility in the realms of medicine and education.

Specifically, sensory-motor skills, essential for interacting with the external environment and executing diverse tasks, bear significant relevance to neurological and mental health assessments. This is especially true for children with learning disabilities, ADHD, and autism spectrum disorders [19, 20]. Processing speed stands as a foundational cognitive element, with ties to age-related cognitive decline [21] and cognitive impairments in psychiatric disorders [22]. Attention, especially sustained attention or vigilance, is pivotal to cognitive processes. Impairments in this domain correlate with conditions like ADHD [23], depression [24], and schizophrenia [25]. Such impairments also relate to real-world challenges, including academic struggles, diminished work performance, and interpersonal relationship difficulties [26]. Executive functions, which encompass cognitive control, planning, and goal setting, are crucial for an individual's adaptability to novel settings, mental health, and overall life quality. Notably, deficits in these functions correlate with neurodevelopmental disorders like ADHD [27], cognitive decline in aging [28], progression from mild cognitive impairment to dementia [29], and assorted emotional disorders [30–32]. Based on the latest model of executive function [33], we focus on two independent components of executive function: cognitive flexibility and working memory, and their related paradigms to cover the measurement of executive function in the most parsimonious way. Spatial ability, the capability to produce, retain, retrieve, and modify structured visual representations, is pertinent for tasks like spatial navigation, map interpretation, and route determination [34]. Deficits in spatial ability have associations with learning disorders, autism, and may serve as a marker for preclinical Alzheimer's disease [35, 36].

Considering the pivotal role of cognitive assessment in both clinical and educational domains, and in light of the growing demand for electronic, standardized, and mobile-based evaluations, this study endeavors to create and validate the Baguan Online Cognitive Assessment System (BOCAS) for both its reliability and validity. BOCAS is designed to fill a gap in cognitive assessment by providing a comprehensive, yet adaptable, test that can meet the diverse needs of both clinical and educational contexts. This paper presents the development of BOCAS, its constituent modules, and the validation procedure, underscoring its potential utility in assessing general cognitive capability (GCA) within the Chinese milieu.

Method

Participants

In this study, a convenience sampling method was employed. The sample size was determined based on the guideline that it should be more than 20 times the number of measurement items (tasks) [37–39]. Participants were openly recruited from campus online forums in Haidian District, Beijing, and are likely to have relatively similar socio-economic backgrounds. The inclusion criteria for participants included: (a) self-reported mental and psychological health, (b) prior experience with touchscreen devices, including smartphones or tablets, and (c) at least a primary school level of education with the ability to independently comprehend task instructions. The exclusion criteria included: (a) a clinical diagnosis of severe mental illnesses such as depression, anxiety, schizophrenia, or bipolar disorder within the past two months, (b) hearing impairment that prevents effective communication at normal conversational volumes, and (c) color blindness or color weakness. After recruitment, we conducted telephone interviews to confirm that participants met the inclusion criteria. Only those who met al.l the inclusion criteria and did not meet any of the exclusion criteria were invited to participate in the experiment. Participants were informed that their participation was voluntary, and they could withdraw at any time. They were also requested to retake the test one month later.

After excluding participants who did not complete the test, 151 healthy adults participated in the initial tablet-based general cognitive assessment. These participants were aged between 18 and 40 years (mean age 22.69 ± 4.60) with an educational background ranging from 7 to 23 years (mean 14.98 ± 2.27). The gender distribution was 30% men and 70% women. For the retest, 124 subjects were included, aged between 18 and 40 years (mean age 22.60 ± 4.54), with educational experiences spanning 7 to 23 years (mean 15.02 ± 2.32). The gender breakdown for this group was 25% men and 75% women. All participants received appropriate compensation upon completion of the experiment. The study secured approval from the Ethics Committee of Beijing Language and Culture University (2023-26).

BOCAS

In the Baguan Online Cognitive Assessment System (BOCAS), the term "Baguan" is derived from an ancient Chinese methodology that uses eight facets to evaluate individuals' talents and competencies. This name was chosen for the current cognitive assessment system to evoke metaphorical connotations relevant to Chinese cultural contexts. Within this linguistic frame, the number eight, represented by the term "Ba" in Pinyin, can be seen as a symbolic figure, suggesting comprehensiveness or totality. BOCAS V1.0, targeting general cognitive ability (GCA), serves as the prototype for the anticipated suite of cognitive assessment offerings under the BOCAS® brand. This suite is envisioned to encompass cognitive evaluations and data management services tailored to diverse specific contexts.

The BOCAS was crafted utilizing the Godot engine (version 3.5.2, available at https://godotengine.org). The software includes several cognitive assessment task modules (described in detail in the next section), with a complete automatic guidance, task interruption redo, and data storage export mechanism. It also contains a local database, with assessment data stored in local devices, allowing for independent export offline and rapid multi-device upload via a local area network.

While the software is compatible with all touch-screen devices, we standardized the device type in this study due to potential influences of device model and screen size on the results. For testing, we utilized the Lenovo Xiaoxin Pad model 13.1.580 (Lenovo TB-J606F) equipped with 6GB of RAM. This tablet features an 11-inch screen and boasts a resolution of 2000×1200.

Procedure

Each participant conducted the assessment in a designated testing space and was provided a specific tablet for the self-directed tasks. The examiner instructed participants to complete the tasks sequentially, and an introductory practice was available for each task. If any unexpected issues arose during the process, the participants could ask the examiner for guidance. While breaks between tasks were permitted, participants were reminded to maintain focus and complete each task independently. Typically, the assessment takes about 30 min to finish, and tasks cannot be revisited. Upon completion, data was saved to a designated folder on the tablet and subsequently organized and analyzed by the examiner.

After the tablet-based assessment, the subjects were required to fill out a self-rated General Cognitive Flexibility questionnaire and the DASS-21 anxiety, depression, and stress questionnaire, which took about 10 min. Then, all the participants undertook the standard version of the Raven's Progressive Matrices test, requiring roughly 20 min. The total duration for this assessment session was approximately 1 h.

A follow-up retest was administered 4 weeks later, mirroring the initial procedure with the exception of the Raven's Progressive Matrices test. First, the tablet tasks were conducted, with the order of tasks and the presentation of stimuli within each task based on a random sequence, ensuring that it did not repeat the sequence from the initial test. This was followed by the completion of self-report questionnaires.

Brief description of tasks

Sensory-motor module (task name: quick tap)

Adapted from a simple target tap task [40], participants are required to quickly and accurately tap falling circular stimuli to assess their basic reaction and visual-motor integration abilities and need to tap as many falling stimuli as possible accurately. Stimuli randomly generate from 5 spawning points at the top of the screen, with an interval of 500ms and a falling speed of 400 screen pixels per second. After a stimulus is tapped, it disappears with a 200ms fading animation. If the stimulus is not tapped before it falls off the bottom of the screen, it is considered as a miss. The task contains 120 trials and lasts about 2 min. Correct hits and misses are accompanied by corresponding auditory feedback. The recorded task data includes the number of correct hits and the center error distance between the actual click position of each button and the optimal hit area in the center. The center error distance is calculated as the relative distance between the click position and the center of the stimulus compared to the size of the stimulus, ranging from 0 to 1. Hitting in the center is record as 0, while a complete miss is record as 1.

Processing speed module (task name: continuous addition)

Derived from the Paced Auditory Serial Addition Test (PASAT) in neurocognitive assessment [41], Participants are required to continuously add the last numbers presented, search for and select the correct answer as quickly as possible, to assess their processing speed and updating ability. Numbers are presented sequentially at a rate of one every 2 s. No response is required for the first number. From the second number onwards, participants need to add it to the previous number. There is a selection number pad from 1 to 20 at the bottom of the panel. The presentation of numbers is randomly controlled according to the answer (2–20), with the answer evenly distributed in the number pad. Subjects need to respond quickly before the next number appears. If the response exceeds the time limit and the next stimulus number appears, it is considered as a miss. Correct, incorrect, and missed responses are accompanied by corresponding auditory feedback. Task data recorded include the reaction time for each response (i.e., the time from the appearance of the number to the response) and the number of correct responses. The task consists of 3 sets of 20 trials.

Sustained attention module (task name: clock judgement)

Based on the Mackworth Clock Continuous Attention Alert Task [42], participants are asked to quickly and accurately judge whether the clock is ticking abnormally to assess their sustained attention function. In this task, the clock moves regularly at a rate of 1 degree per second, but at certain time it will suddenly speed up to 2 degrees per second, forming a type of leap-second movement. Participant need to quickly respond to the abnormal movement by pressing a button. It is considered a hit if the response was made within 1s. In contrast, if an abnormality is not detected or the response was longer than 1s, it is considered a miss. If the key is pressed when there is no abnormal movement, it is incorrect (false alarm). This task lasts for 6 min, with the probability of abnormal movement appearing 4–6 times per minute. Incorrect responses are accompanied by corresponding auditory feedback, and there is no feedback reminder when an abnormal movement is missed. The recorded data includes the number of correct hits, missed hits, and incorrect hits, and the response time for each instance.

Working memory modulate (task name: memory blocks)

Adapted from the Corsi Spatial Working Memory task [43], participants is asked to recall the positions of blocks in sequence to test their working memory capacity. The maximum number of square layouts is 9 with random position, which will be lit for 0.2 s in yellow. The order of the lighting is the sequence to be recalled. After the prompt "Please Recall" is shown, subjects are required to tap on the blocks that were just lit in order. If the response is correct, the block turns green; if incorrect, it turns red. The number of blocks to be recalled starts from two and gradually increases until it stops after two consecutive errors. If it does not exceed 9 blocks, the memory capacity is recorded as the maximum number of remembered blocks when stopped. The test comprises 2 sets, with the final memory capacity being the average of the two. It should be noted that compared with the design of fixed format (such as 3*3), in the current design, the layout of the blocks for each memory trial will change to increase memory load. The data recorded include the response position of each recall, correct and incorrect responses, and the highest number of items remembered.

Cognitive flexibility module (task name: trail making test)

It is an adapted version of the paper-and-pencil Trail Making Task [44]. In this task, participants are required to alternately click on the number and dot buttons in ascending order to evaluate cognitive flexibility and integrated response capabilities. Participants need to quickly search and click on a randomly generated panel of numbers and dots, following the order of "number 1" - "dot 1" - "number 2" - "dot 2". "number 9" - "dot 9". If the participant clicks correctly, the number or dot button will turn from gray to blue, if the participant clicks incorrectly, the button will not turn blue and cannot proceed to the next click. The task contains a total of 3 sets, the layout is randomly distributed. Correct and incorrect responses are accompanied by corresponding visual feedback. The data recorded by the task is the reaction time for correctly completing each set.

Spatial ability module (task name: spatial navigation)

This task has been modified from a spatial navigation task [45], in which participants are required to search for a visual target in a virtual circular environment that was hidden after spatial rotation with different reference clues. Specifically, there will be a rotating circle on the screen. There is a target object, a small ball, in a certain position in the circle. Participants need to remember the specific position of the small ball. After pressing the "Memorized" button, the circle rotates and the small ball disappears. There is a starting response point on the circle that represents the test subject's perspective. The test subject needs to drag a route from this point to find the position before the small ball disappeared. The task is divided into three stages. In the first stage, self+environmental navigation is required, the starting response point does not change before and after rotation (i.e., self as the reference), and there are two tree-shaped references (i.e., environmental reference). In the second stage, self-reference navigation is required, the starting response point does not change before and after rotation, and there are no reference objects. In the third stage, environmental navigation is required, the starting response point changes before and after rotation, and there are two tree-shaped reference objects. The position of the target object, the small ball, is pseudorandom. In each stage, seven trials will be taken, that is, the circle will rotate seven times, and the rotation angle is also pseudorandom (45°, 90°, 135°, 180°, 225°, 270°, 315°). Participants need to respond within the circle area, the response time is not limited, and the answer can be changed. The results recorded include the error distance between the confirmed target position and the real target position each time. The schematic diagram of the six modules can be seen in Fig. 1.

Data analysis

The exported data format for the task modules is .CSV. Custom Python scripts are used to organize the exported local data, including aligning the data according to the test subject number, cleaning and conducting preliminary analysis on the raw data, merging demographic data and the scores of the subjects' subjective cognitive assessment scales and Raven's intelligence test.

For specific task performance scores, the Sensory-Motor (Quick Tap) task collects the number of correct clicks and the center error. The calculated index for this task=the number of correct clicks / total number * (1 – mean (center error)). The Processing Speed (Continuous Addition) task records information such as correct, missed, and erroneous reactions. The calculated index for this task=the average number of correct responses in the Continuous Addition task / total number of required responses. The Sustained Attention (Clock Judgment) task records correct, erroneous, and missed responses. The calculated index for this task=number of correct responses - number of erroneous responses. The Working Memory (Memory Blocks) task records correct and erroneous recall responses. The calculated index for this task=the average maximum number of blocks recalled. The Cognitive Flexibility (Trail Making Test) task records the reaction time of each response. The calculated index for this task=average time to complete the Trail Making Test (only when the current click is correct can the next button be activated and become clickable, so the completion time can well reflect individual ability). The Spatial Ability (Spatial Navigation) task records the error distance of each response. The calculated index for this task=average error distance in the Spatial Navigation Task.

It should be noted that indices calculated based on hit rates, number of responses, etc., are positively correlated with actual task performance, while those based on average error distances and average reaction times are inversely related to actual task performance and function.

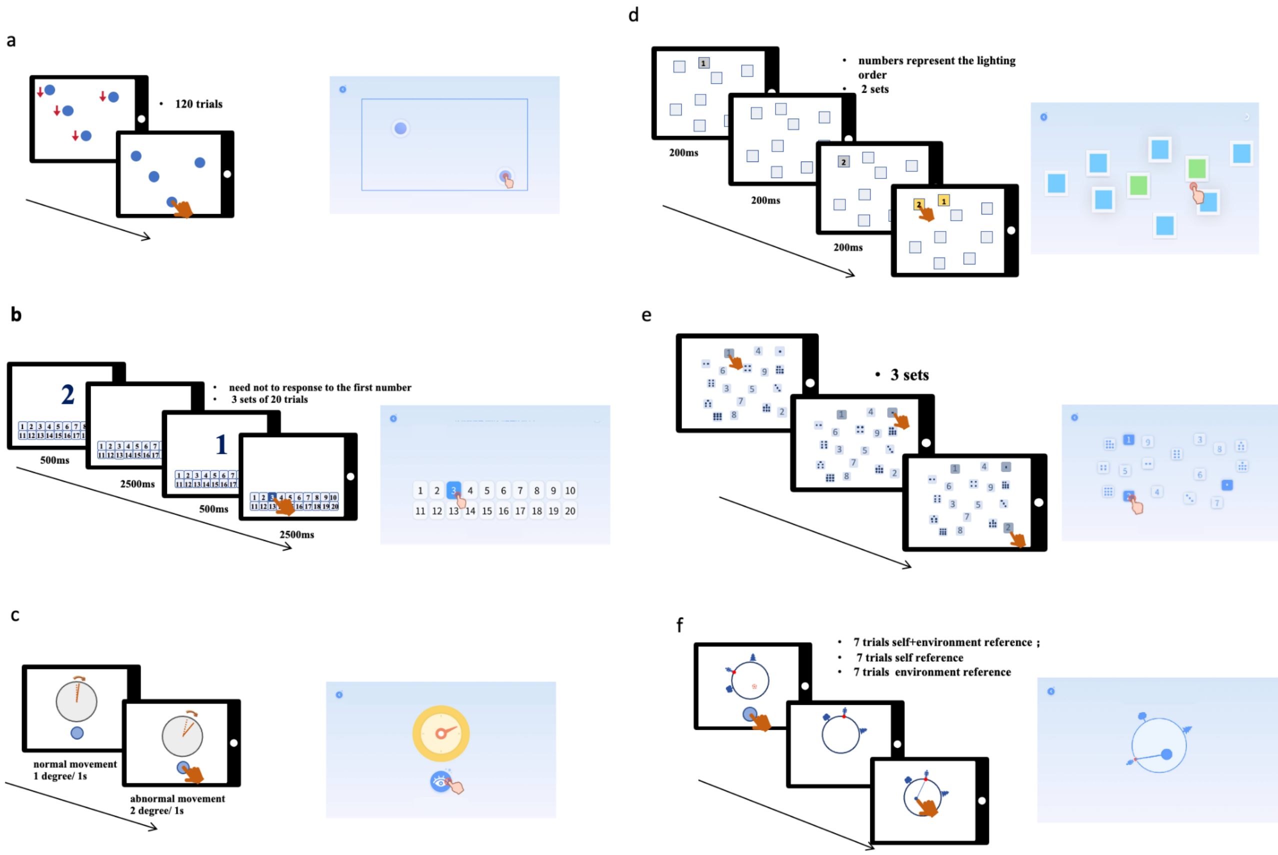

Fig. 1 Illustration of the six modules in the tablet-based general cognitive ability test. **a**-**f** represent the diagrams of the six modules of sensory-motor module, processing speed module, attention module, working memory module, cognitive flexibility module, and spatial ability module respectively. In each subgraph, the right side is the sequence diagram, and the left side is the real screenshot

The calculated indices are further transformed into standard scores for each module's performance through the formula 100±15*Z for further analysis. In addition, we used two strategies to calculate the general cognitive ability score (GCA) for the test set: (a) the standard scores of each module are simply summed to form the assessment total score; (b) factor score based on principal component analysis was obtained to get the weighted measurements of GCA. The study's reliability analysis uses internal consistency (Cronbach's Alpha), split-half, and test-retest reliability using Pearson's correlation coefficient as well as Intraclass Correlation Coefficient (ICC). Item analysis uses item-total correlation and independent sample t-tests for high and low groups. Construct validity testing uses exploratory factor analysis for pre-test data and confirmatory factor analysis for post-test data. Taking the scores of the Raven's Reasoning Test and the subjective cognitive flexibility evaluation as the criterion, Pearson correlation analysis was used to examine the criterion validity of each independent task and the overall performance of the test. In addition, this study also explored the relationship between task performance and demographic information using independent and paired sample t-tests, correlation analysis. Data analysis is performed using SPSS.22 and RStudio with related packages such as Lavaan, semPlot.

Result

Reliability analysis

Reliability denotes the consistency of results when a method is repeatedly applied to measure the same entity. In behavioral assessments, the internal consistency of multiple trials within a test can be represented using the internal consistency coefficient (Cronbach's Alpha). Additionally, the consistency of test scores across different time points, known as test-retest reliability, further illustrates this attribute. Table 1 displays the internal consistency coefficients for the six task modules alongside the correlation of test performance one month apart (test-retest reliability using Pearson's correlation coefficient). The table also presents the calculated indices for each module, which serve as benchmarks for determining test-retest reliability. The results indicate that the internal consistency coefficients for each BOCAS dimension exceed 0.7, with test-retest reliability (both r and ICC) ranging from 0.56 to 0.71, all exceed 0.5 and significant

Table 1 Reliability analysis of BOCAS

Modules	Cronbach's Alpha ($n=151$)	Internal repetitions	Calculated score for each module	Time 1	Time 2	Test-retest r ($n=120$)	ICC (2,1) 95% CI	T(p)
Sensory-Motor	0.731	2	The Number of Correct Clicks / Total Number * (1 - Center Error)	0.569 (0.08)	0.568(0.07)	0.576***	0.575 (0.445 – 0.682)	0.057 (0.955)
Processing Speed	0.846	3	The Average Number of Correct Responses	12.32(3.69)	14.37(3.68)	0.610***	0.612 (0.489 –0.711)	-7.01 (0.001)
Sustained Attention	0.822	2	Number Of Correct Hits - Number of Erroneous Hits	13.61(7.84)	13.37(7.11)	0.708***	0.708 (0.608 – 786)	0.441 (0.660)
Working Memory	0.712	4	The Average Maximum Number of Memory Blocks Recalled	5.99(0.99)	6.13(1.05)	0.602***	0.604 (0.478 –0.705)	-1.65 (0.100)
Cognitive Flexibility	0.757	3	Average Reaction Time	17.799(4.04)	17.365(3.25)	0.562***	0.563 (0.420 –0.673)	1.411 (0.161)
Spatial Ability	0.776	3	Average Error Distance	0.1282(0.07)	0.1261(0.06)	0.623***	0.625 (0.505 –0.722)	0.420 (0.675)
Total score of BOCAS	0.651	6	Sum of standard scores for each module	601.67(54.58)	599.76(56.02)	0.782***	0.783 (0.704 –0.843)	0.579 (0.542)

Note: The Sensory-Motor module and the Sustained Attention module are split into odd and even halves for the process response, from which the split-half reliability is extracted. For other tasks with internally set repetitions, the internal consistency reliability of the repeated tests is reported. Here, the test-retest reliability is reported using Pearson's correlation coefficient r, as well as the Intraclass Correlation Coefficient (ICC) value for Single Measures under the two-way random, absolute agreement model, ICC (2,1). ***: $p<0.001$

at the 0.001 level, indicating moderate to good test-retest reliability [46, 47].

We further conducted a difference analysis on each of modules before and after a 4-week interval to investigate the practice effect. Descriptive data on the indicators of each module in the two measurement times, as well as pairwise differences, are presented in Table 1. The results showed that only Processing Speed showed a significantly better performance in the second measurement. For other task modules, the practice effect was not significant. Furthermore, the results show that if the standard score of each module is used as the consistency measurement index, the overall internal alpha coefficient of the tool is 0.651; the test-retest reliability of the GCA can reach 0.782 for total score and 0.775 for factor score; while there was no significant difference in the BOCAS total score in the measurements at the two time points (t, p), which indicated the cross-time stability of the evaluation. Related reports are shown in Table 1.

Modules analysis

The test is designed to assess an individual's general cognitive function. Each module and its measurement dimensions are derived from theoretical deduction. It is necessary to evaluate the correlation between the performance of a single module and the individual's comprehensive cognitive function, as well as the discrimination between each dimension. The correlation between the scores of each dimension and the total score, that is, the item-total correlation, is calculated. Here we calculated the correlation between a certain task and the total score without that particular task to examine the contribution of a certain task to the overall score. The results show that the correlation between the task score and the total scores of other tasks reached a significance of $p<0.01$. Demonstrated good internal consistency and convergency of the test. The top 27% of the total scores are classified as the high-score group and the bottom 27% as the low-score group, with 40 people in each group. The independent samples t-test shows that there are significant differences between the high and low-score groups in each module dimension, that is the discrimination of the BOCAS module meets the basic requirements. The related results are shown in Table 2.

Table 2 Discrimination analysis of each module of BOCAS

Modules	Task-total correlation (The task excluded)	t-value for high and low score groups
Sensory-Motor	0.273**	5.637***
Processing Speed	0.400***	11.098***
Sustained Attention	0.342***	3.064**
Working Memory	0.355***	7.717***
Cognitive Flexibility	0.519***	8.145***
Spatial Ability	0.398***	6.164***

Note: **: $p<0.01$, ***: $p<0.001$

Validity test

Structural validity

We used the exploratory factor analysis to test the structural validity of BOCAS. The KMO value in this study is 0.749, so the factor analysis can be performed (generally, if the KMO value is greater than 0.5, a factor analysis can be conducted). The Bartlett's test, $\chi^2=99.42$, $p<0.001$,

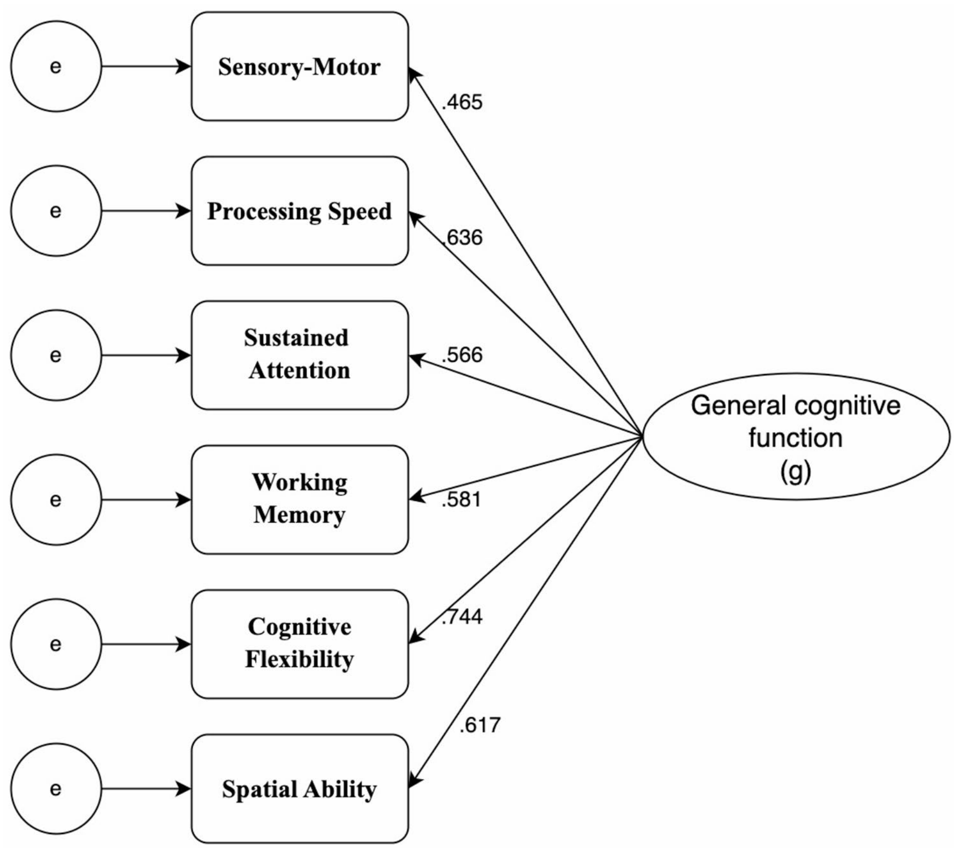

Fig. 2 Factor loading plot for each module of the BOCAS

Table 3 Results of confirmatory factor analysis

Model	χ2/df	CFI	TLI	RMSEA
Single Factor Model	1.81	0.932	0.887	0.080

Note: Chi-Square (χ^2) and Chi-Square to Degrees of Freedom Ratio (**χ^2/df**): A χ^2/df ratio less than 3 is generally considered acceptable, and a value less than 2 is preferred; Comparative Fit Index (**CFI**): A CFI value greater than 0.90 is generally considered acceptable, and a value greater than 0.95 is preferred; Tucker-Lewis Index (**TLI**): A TLI value greater than 0.85 is generally considered acceptable, and a value greater than 0.90 is preferred; Root Mean Square Error of Approximation (**RMSEA**): An RMSEA value less than 0.05 indicates a close fit, values between 0.05 and 0.08 indicate a reasonable fit, and values greater than 0.10 indicate a poor fit

suggests the existence of common factors. By using the principal component analysis method and extracting common factors with eigenvalues greater than 1, and then rotating using the Varimax method, one factor was extracted, which can explain 36.92% of the total variance, in line with the general cognitive g factor conception. For the post-test sample, the fixed one factor model can explain 39.60% of the total variance. At the same time, confirmatory factor analysis was used to construct a single-factor model for the post-test data. The maximum likelihood estimation method was used to test the fit degree of the data and the model. χ^2/df is less than 2, CFI and TLI are both greater than 0.85, and RMSEA equals 0.080, indicating that the single-factor model fits the observed data quite well [48–51]. Figure 2; Table 3 shows how each module contributes to the overall general cognitive ability and related fitting coefficients.

Criterion-related validity

The study obtained the participants' scores from the Standard Raven's Progressive Matrices test as the criterion scores for general cognitive abilities and conducted a correlation analysis with the BOCAS scores. The results showed a significant correlation between the GCA score of BOCAS and the score of the Raven's Progressive Matrices test (Total score: $r=0.58$, $p<0.001$; factor score: $r=0.59$, $p<0.001$). The scores of each test module were also significantly correlated with the Raven's score, specifically: Sensory-Motor ($r=0.27$, $p=0.001$), Processing Speed ($r=0.41$, $p<0.001$), Sustained Attention ($r=0.32$, $p<0.001$), Working Memory ($r=0.28$, $p=0.01$), Cognitive Flexibility ($r=0.58$, $p<0.001$), and Spatial Ability ($r=0.38$, $p<0.001$). These results indicate that BOCAS has good criterion-related validity.

Moreover, age-related developmental changes can also act as a criterion for validating cognitive tests. The GCA score of BOCAS is significantly negatively correlated with age (Total score: r = -0.32, $p<0.001$; Factor score: r = -0.33, $p<0.001$). In terms of the individual modules, the scores of Processing Speed (r = -0.26, p=0.002), Working Memory (r = -0.19, p=0.020), Cognitive Flexibility (r = -0.22, p=0.007), and Spatial Ability (r = -0.44, p<0.001) are significantly negatively correlated with age; Sustained Attention and Sensory-Motor are not modulated by age.

Furthermore, the study obtained the participants' self-reported cognitive flexibility scores as a subjective measure of cognitive ability and conducted a correlation analysis with the BOCAS scores. The results showed a significant correlation between the GCA of BOCAS and the score of subjective cognitive flexibility (Total score: r=0.35, p<0.001; Factor score: r=0.352, p<0.001). In terms of the individual modules, the scores of Processing Speed (r=0.38, p<0.001), Cognitive Flexibility (r=0.21, p=0.004), Sustained Attention (r=0.24, p=0.004), and Sensory-Motor (r=0.18, p=0.029) were significantly positively correlated with cognitive flexibility; Spatial Ability and Working Memory were not modulated by the subjective cognitive flexibility score. Table 4 describes the correlation coefficients of BOCAS and its various modules with Raven's score, age, and self-reported cognitive flexibility.

Preliminary application of BOCAS

In order to better examine the relationship between demographic variables (such as gender and education level, etc.), subjectively reported emotional states (i.e., DASS score) and test performance, we firstly performed a mixed-ANOVA based on gender (male/female) and two time points (time 1/time 2) for each module and the GCA score. The results only showed the main effect over time in the Processing Speed (as reported above), and no other gender-related main effect or interaction effects were found (p>0.282).

Furthermore, we conducted a partial correlation analysis on the GCA score of BOCAS and separated modules with individuals' level of education controlling for age. The results indicated a positive correlation between the GCA scores of BOCAS and the modules of Sensory-Motor and Cognitive Flexibility in the pretest and the Spatial Ability in both two test times, suggesting an individual's education level can affect his/her cognitive ability performance (p_s< 0.05).

In addition, the analysis of BOCAS scores and individual emotional status found that there is a correlation between task-based cognitive function and individual's emotional status. The DASS score has a significant negative correlation with the GCA score of BOCAS (p_s<0.005), which is consistent with the hypothesis that individuals at risk of emotional disorders are related to cognitive deficits. However, using the Raven's Progressive Matrices test was unable to sensitively show this association, indicating that this research tool has the potential to be applied in cognitive evaluation in the cross-diagnosis of mental illnesses. Relevant data are presented in Table 5.

Table 4 Criterion-related validity of the BOCAS GCA score and its individual modules

	Raven's score(r)	age(-r)	Score of subjective cognitive flexibility
Total score of BOCAS	0.583***	−0.321***	0.354***
Factor score of BOCAS	0.587***	−0.328***	0.352***
Sensory-Motor	0.268***	−0.117	0.181*
Processing Speed	0.410***	−0.256**	0.375***
Sustained Attention	0.324***	−0.049	0.238**
Working Memory	0.281**	−0.190*	0.157
Cognitive Flexibility	0.583***	−0.217**	0.211**
Spatial Ability	0.383***	−0.438***	0.136

Note: **: p<0.01, ***: p<0.001

Discussion

This study sought to fulfill the growing demand for a dependable and flexible cognitive assessment instrument by introducing the Baguan Online Cognitive Assessment System (BOCAS). BOCAS provides a tablet-centric assessment suite tailored to gauge an individual's general cognitive prowess across multiple cognitive domains. Through rigorous validation, BOCAS has showcased significant promise for both clinical and educational applications.

Reliability stands as a cornerstone for any assessment instrument. The internal consistency reliabilities (α) for each of the six modules within BOCAS span from 0.72 to 0.85. These figures denote a moderate to high level of internal consistency, implying that the modules provide consistent and dependable outcomes. Furthermore, the test-retest reliabilities of these modules, which range from 0.56 to 0.71, demonstrate substantial temporal stability [52]. Such results are significant, especially when considering that the typical test-retest reliability for cognitive behavioral tasks averages around 0.61 [53]. Of paramount importance is the observation that the test-retest reliability of BOCAS's GCA score approaches 0.78, even surpassing the common benchmarks set by self-report measures [53]. The test-retest reliability of BOCAS's modules, exceeding these standards, emphatically underscores the tool's robust consistency and reliability.

A standout feature of BOCAS is its proficiency in countering the practice effect, a prevalent concern in cognitive assessment. Practice effects have the potential to skew assessment interpretations, as individuals might exhibit improved performance merely due to prior test exposure. Yet, BOCAS effectively curtails this influence, as illustrated by the paired-sample t-tests that revealed negligible performance variances between initial and subsequent tests. This holds significant merit

Table 5 Task performance in each module and gender, education and emotion

	Time	Sensory-Motor		Processing Speed		Sustained Attention		Working Memory		Cognitive Flexibility		Spatial Ability		GCA (Total)		GCA (Factor)	
		F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M
Gender (Score)	T1	101.75	95.73	101.29	96.85	100.42	98.96	99.94	100.14	101.51	96.32	100.55	98.64	605.19	586.66	0.10	-0.24
	T2	100.69	97.92	102.10	93.69	100.05	99.82	100.55	98.33	101.10	96.69	100.81	97.54	605.32	584.02	0.10	-0.30
Education (pr)	T1	0.213*		0.144		0.101		0.163		0.223*		0.411***		0.341***		0.339***	
	T2	0.147		0.153		0.108		0.121		0.148		0.253**		0.255**		0.244**	
DASS (r)	T1	-0.029		-0.113		-0.276***		-0.123		-0.241**		-0.103		-0.230**		-0.238**	
	T2	-0.100		-0.065		-0.215*		-0.027		-0.219*		-0.174		-0.213*		-0.203*	

Note: *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$

since practice effects can inject biases into assessment outcomes [54, 55]. Practically, by diminishing the practice effect, BOCAS bolsters its capability to accurately track cognitive shifts over durations, be it for monitoring developmental trajectories or gauging the efficacy of interventions in both clinical and academic contexts.

The capacity for discrimination — the ability of an assessment tool to differentiate between high and low performers — is pivotal [56]. Every module within BOCAS showcased outstanding discriminatory power. A comparative analysis of scores between the top and bottom 27% of participants for each module revealed pronounced differences. Such results underscore BOCAS's adeptness in detecting nuanced disparities in cognitive performance across individuals. Beyond its reliability, this robust discrimination underlines BOCAS's precision in evaluating cognitive competencies.

A foundational principle of cognitive assessment posits the existence of a general cognitive ability factor (g factor) that underpins performance across diverse cognitive tasks. Approximately 40% of the total variation in scores from a test set can be ascribed to this single factor [57]. In alignment with this principle, exploratory factor analysis (EFA) on BOCAS identified a singular factor responsible for roughly 40% of the total variance. Each module demonstrated significant loadings on this general cognitive ability component. Subsequent confirmatory factor analysis (CFA) further endorsed the single-factor model, showcasing optimal fit indices. These results underscore the capacity of BOCAS's six modules to effectively measure general cognitive ability.

Criterion-related validity is a vital aspect of any cognitive assessment tool. BOCAS showcased robust criterion-related validity through its significant correlation with the standard Raven's Progressive Matrices, a widely recognized measure of nonverbal fluid intelligence [58]. Importantly, both the General Cognitive Ability (GCA) from BOCAS and the performance across individual modules correlated markedly with the Raven's scores. This emphasizes BOCAS's competence in gauging general cognitive ability, even without explicit reasoning tasks. The modules encompass spatial memory, navigation, attention, cognitive processing, and flexibility, ensuring that each test distinctly captures individual cognitive functions while collectively offering a holistic view of an individual's cognitive prowess. The pronounced correlation coefficient for the GCA score posits that BOCAS might hold an edge over traditional single-form fluid IQ assessments.

Differences in cognitive performance across age groups serve as a pivotal criterion for the validation of cognitive assessments. Many cognitive abilities tend to peak in early adulthood and subsequently decline with advancing age [59]. BOCAS effectively captured these age-related variations in performance, with older individuals generally showcasing diminished cognitive outcomes, further bolstering its credibility as a cognitive assessment instrument. Moreover, when age was accounted for, educational attainment displayed a robust positive correlation with performance on BOCAS. This aligns with the well-established notion that education has a beneficial impact on cognitive function [60]. Such findings highlight BOCAS's adeptness in discerning cognitive disparities influenced by both age and education, thereby reinforcing its legitimacy as an evaluative tool.

Beyond its established psychometric strengths, BOCAS stands out due to its adaptability and broad range of potential uses. The significant correlations between BOCAS scores and self-reported cognitive flexibility, paired with associations with adverse emotions such as depression and anxiety, underscore its potential to illuminate the cognitive facets underlying emotional well-being. Specifically, emotional disorders involve solidified irrational beliefs and pathological rumination, which are closely related to poor cognitive flexibility and the inability to update and inhibit emotional interference [61]. Cognitive flexibility deficits and diminished attentional control could underlie symptom emergence and be crucial to their perpetuation [62]. Our data revealed that besides the overarching GCA score, BOCAS's attention and cognitive flexibility module performances negatively correlated with chronic emotional dysregulation. This positions BOCAS as a potent evaluative instrument for cognitive dynamics linked to emotional disorders [63].

The inclusion of multiple modules representing various cognitive functions allows BOCAS to offer a transdiagnostic perspective, making it invaluable for diverse research and clinical applications. For instance, through its integration of multiple modules representing various cognitive functions such as memory, attention, and executive function, BOCAS offers a holistic view of a patient's cognitive profile, which is crucial for diagnosing and monitoring neurocognitive disorders like Alzheimer's disease (AD) and mild cognitive impairment (MCI) [64]. The ability of BOCAS to conduct rapid screenings in primary care settings allows for the early identification of potential cognitive issues, enabling timely interventions that can significantly delay the progression of these conditions [65]. Based on the detailed data provided by BOCAS, healthcare teams can tailor rehabilitation programs to suit individual patient needs, enhancing treatment efficacy and patient engagement, while regular follow-ups using the tool enable the tracking of progress and adjustment of therapeutic strategies. The digital nature of BOCAS also supports the expansion of telemedicine services, making cognitive assessments more accessible to those in remote areas or with mobility challenges [66]. Additionally, the large datasets collected through BOCAS contribute to ongoing research, aiding in the discovery of new mechanisms behind cognitive disorders, the development of novel treatments, and the improvement of existing assessment tools [67], thereby creating a positive feedback loop between research and clinical practice.

Limitations

While the BOCAS shows significant promise, several limitations should be acknowledged. Firstly, the study sample, which was recruited through social and convenience sampling, lacked a systematic distribution of demographic and socioeconomic variables. This imbalance may limit the generalizability of the results to broader and more diverse populations. Additionally, the validity and reliability of BOCAS have not been tested in clinical populations, which is a critical gap that needs to be addressed in future research.

Another limitation is that BOCAS does not cover some broad cognitive domains, such as social and emotional cognition. These areas are increasingly recognized as important components of overall cognitive function, and their inclusion would enhance the comprehensiveness of the assessment. Furthermore, more research is needed to monitor long-term cognitive changes, as the current study primarily focuses on cross-sectional data. Longitudinal studies will provide valuable insights into the system's ability to track cognitive development and decline over time.

Accessibility is another concern, particularly for individuals with limited access to technology. Ensuring that BOCAS is accessible to a wider range of users, including those in underserved communities, is essential for its widespread adoption and utility. Additionally, the study was limited to Chinese adults, and further research is needed to validate the system's effectiveness and reliability in different cultural and linguistic contexts.

Despite these limitations, BOCAS in its current form is promising and has the potential for wide application in various domains. Future research with larger and more diverse sample groups, including clinical populations, will further reinforce the overall validity and reliability of the system. Expanding BOCAS to cover additional cognitive domains, such as social and emotional cognition, and improving its long-term monitoring capabilities will make it a more comprehensive and useful tool. The system's flexibility in task module configuration, remote testing capabilities, and robust data management systems will facilitate its expansion into various assessment and application areas, thereby fostering advancements in digital education and healthcare.

Conclusion

As an electronic assessment tool for general cognition, BOCAS demonstrates commendable reliability and validity. Its modular structure and streamlined data collection make it highly adaptable to diverse settings. BOCAS holds great promise in advancing cognitive research, clinical diagnostics, educational assessments, and human resources evaluations. Its potential to enhance our understanding of cognitive functions, diseases, personality traits, and overall performance positions it as a pivotal tool across various fields.

Acknowledgements
We thank the research assistants Dingyao Fan, Tianjiao Wei, Wen Wang, Yuanyuan Shi for their support during the data collection process. We thank Wang Daotuan for providing technical support during software development and data analysis.

Author contributions
D.N.P. and G.X. designed the system; D.N.P. wrote the main manuscript text and conduct data analysis; H.L.X, Y.J.Z, Y.X.Z, C.Z.L, X. M, Y.Y.J, J.J.R, Y.Y.W, performed the testing process; W.W. provides support for data analysis; resources are provided by G.X. All authors reviewed the manuscript.

Funding
This research project is supported by Science Foundation of Beijing Language and Culture University (supported by "the Fundamental Research Funds for the Central Universities") (24QN28, YYBB29) for Dong-ni Pan. We also thank the Hangzhou Biomedical and Health Industry Development Support Science and Technology Project (2022WJC265 to W.W) for its support.

Data availability
The datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate
All procedures performed in the study involving human participants were in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. All experimental protocols were approved by the ethics committee at the Ethics Committee of Beijing Language and Culture University (2023-26). Informed consent was obtained from all individual participants included in the study. The research team placed a high emphasis on ensuring participant privacy and confidentiality, ensuring that all obtained data was exclusively utilized for research objectives.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 11 December 2023 / Accepted: 10 December 2024
Published online: 24 December 2024

References

1. Harvey PD. Domains of cognition and their assessment. Dialogues Clin Neurosci. 2019;21(3):227–37.
2. Spearman C. General intelligence,' objectively determined and measured. Am J Psychol. 1904;15(2):201–93.
3. Carroll JB. Human cognitive abilities: A survey of factor-analytic studies. New York, NY, US: Cambridge University Press; 1993.
4. Kato AE, Scherbaum CA. Exploring the Relationship between Cognitive Ability Tilt and Job Performance. J Intell 2023, 11(3).
5. Morgeson FP, Delaney-Klinger K, Hemingway MA. The importance of job autonomy, cognitive ability, and job-related skill for predicting role breadth and job performance. J Appl Psychol. 2005;90(2):399–406.
6. Kuncel NR, Ones DS, Sackett PR. Individual differences as predictors of work, educational, and broad life outcomes. Pers Indiv Differ. 2010;49(4):331–U336. https://www.sciencedirect.com/science/article/pii/S0191886910001765.
7. Tangalos EG, Petersen RC. Mild Cognitive Impairment in Geriatrics. Clin Geriatr Med. 2018;34(4):563–89.
8. Culpepper L, Lam RW, McIntyre RS. Cognitive Impairment in Patients With Depression: Awareness, Assessment, and Management. J Clin Psychiatry. 2017;78(9):1383–94.
9. Gebreegziabhere Y, Habatmu K, Mihretu A, Cella M, Alem A. Cognitive impairment in people with schizophrenia: an umbrella review. Eur Arch Psychiatry Clin Neurosci. 2022;272(7):1139–55.
10. Sharma SR, Gonda X, Tarazi FI. Autism Spectrum Disorder: Classification, diagnosis and therapy. Pharmacol Ther. 2018;190:91–104.
11. Rubia K. Cognitive Neuroscience of Attention Deficit Hyperactivity Disorder (ADHD) and Its Clinical Translation. Front Hum Neurosci. 2018;12:100.
12. Chen K-L, Xu Y, Chu A-Q, Ding D, Liang X-N, Nasreddine ZS, Dong Q, Hong Z, Zhao Q-H, Guo Q-H. Validation of the Chinese Version of Montreal Cognitive Assessment Basic for Screening Mild Cognitive Impairment. J Am Geriatr Soc. 2016;64(12):e285–90.
13. Sweet L, Adel MV, Metcalf V, Wright L, Harley A, Leiva R, Taler V. The Montreal Cognitive Assessment (MoCA) in geriatric rehabilitation: psychometric properties and association with rehabilitation outcomes. Int Psychogeriatr. 2011;23(10):1582–91.
14. Khaw J, Subramaniam P, Abd Aziz NA, Ali Raymond A, Wan Zaidi WA, Ghazali SE. Current Update on the Clinical Utility of MMSE and MoCA for Stroke Patients in Asia: A Systematic Review. Int J Environ Res Public Health 2021, 18(17).
15. Shuttleworth-Edwards AB. Generally representative is representative of none: commentary on the pitfalls of IQ test standardization in multicultural settings. Clin Neuropsychol. 2016;30(7):975–98.
16. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, Cummings JL, Chertkow H. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc. 2005;53(4):695–9.
17. Maruff P, Thomas E, Cysique L, Brew B, Collie A, Snyder P, Pietrzak RH. Validity of the CogState brief battery: relationship to standardized tests and sensitivity to cognitive impairment in mild traumatic brain injury, schizophrenia, and AIDS dementia complex. Archives Clin Neuropsychology: Official J Natl Acad Neuropsychologists. 2009;24(2):165–78.
18. Sternberg D, Ballard K, Hardy J, Katz B, Doraiswamy PM, Scanlon M. The largest human cognitive performance dataset reveals insights into the effects of lifestyle factors and aging. *Frontiers in Human Neuroscience* 2013, 7%U https://www.frontiersin.org/articles/https://doi.org/10.3389/fnhum.2013.00292
19. Carsone B, Green K, Torrence W, Henry B. Systematic Review of Visual Motor Integration in Children with Developmental Disabilities. *Occupational Therapy International* 2021, 2021:1801196.
20. Kulp MT. Relationship between Visual Motor Integration Skill and Academic Performance in Kindergarten through Third Grade. Optom Vis Sci. 1999;76(3):159U. https://journals.lww.com/optvissci/abstract/1999/03000/relationship_between_visual_motor_integration.03015.aspx.
21. Kail R, Salthouse TA. Processing speed as a mental capacity. Acta Psychol. 1994;86(2–3):199–225.
22. Knowles EEM, David AS, Reichenberg A. Processing speed deficits in schizophrenia: reexamining the evidence. Am J Psychiatry. 2010;167(7):828–35.
23. Barkley RA. Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. Psychol Bull. 1997;121(1):65–94.
24. Hammar A, Ardal G. Cognitive functioning in major depression–a summary. Front Hum Neurosci. 2009;3:26.
25. Carment L, Dupin L, Guedj L, Térémetz M, Krebs M-O, Cuenca M, Maier MA, Amado I, Lindberg PG. Impaired attentional modulation of sensorimotor control and cortical excitability in schizophrenia. Brain. 2019;142(7):2149–64.
26. Nigg JT. Attention-deficit/hyperactivity disorder and adverse health outcomes. Clin Psychol Rev. 2013;33(2):215–U228. https://www.sciencedirect.com/science/article/pii/S0272735812001705.
27. Seidman LJ. Neuropsychological functioning in people with ADHD across the lifespan. Clin Psychol Rev. 2006;26(4):466–85.

28. De Luca CR, Wood SJ, Anderson V, Buchanan J-A, Proffitt TM, Mahony K, Pantelis C. Normative data from the CANTAB. I: development of executive function over the lifespan. J Clin Exp Neuropsychol. 2003;25(2):242–54.
29. Junquera A, García-Zamora E, Olazarán J, Parra MA, Fernández-Guinea S. Role of Executive Functions in the Conversion from Mild Cognitive Impairment to Dementia. J Alzheimer's disease: JAD. 2020;77(2):641–53.
30. Chai WJ, Abd Hamid AI, Abdullah JM. Working Memory From the Psychological and Neurosciences Perspectives: A Review. Front Psychol. 2018;9:401.
31. Kirova AM, Bays RB, Lagalwar S. Working memory and executive function decline across normal aging, mild cognitive impairment, and Alzheimer's disease. *Biomed Res Int* 2015, 2015:748212.
32. Martinussen R, Hayden J, Hogg-Johnson S, Tannock R. A Meta-Analysis of Working Memory Impairments in Children With Attention-Deficit/Hyperactivity Disorder. J Am Acad Child Adolesc Psychiatry. 2005;44:377–84.
33. Feng J, Zhang L, Chen C, Sheng J, Ye Z, Feng K, Liu J, Cai Y, Zhu B, Yu Z, et al. A cognitive neurogenetic approach to uncovering the structure of executive functions. Nat Commun. 2022;13(1):4588.
34. Lohman DF. Spatial ability and g. Human abilities: Their nature and measurement. edn. Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc; 1996. 97–116%@ 978-110-8058-1800-8050.
35. Coughlan G, Laczó J, Hort J, Minihane A-M, Hornberger M. Spatial navigation deficits - overlooked cognitive marker for preclinical Alzheimer disease? Nat Reviews Neurol. 2018;14(8):496–506.
36. Fragueiro A, Tosoni A, Frisoni M, Di Matteo R, Sestieri C, Committeri G. Travel in the Physical and Mental Space: A Behavioral Assessment of the Phylogenetic Continuity Hypothesis Between Egocentric Navigation and Episodic Memory. Evol Psychol. 2021;19(3):14747049211040823.
37. Bujang MA. A Step-by-Step Process on Sample Size Determination for Medical Research. Malays J Med Sci. 2021;28(2):15–27.
38. Burmeister E, Aitken LM. Sample size: How many is enough? Australian Crit Care. 2012;25(4):271–4.
39. Osborne J, Costello AB. Sample size and subject to item ratio in principal components analysis. Prac Assess Res Eval 2004, 9.
40. Brenner E, de la Malla C, Smeets JBJ. Tapping on a target: dealing with uncertainty about its position and motion. Exp Brain Res. 2023;241(1):81–104.
41. Tombaugh TN. A comprehensive review of the Paced Auditory Serial Addition Test (PASAT). Archives Clin Neuropsychology: Official J Natl Acad Neuropsychologists. 2006;21(1):53–76.
42. Giambra LM, Quilter RE. Sustained attention in adulthood: a unique, large-sample, longitudinal and multicohort analysis using the Mackworth Clock-Test. Psychol Aging. 1988;3(1):75–83.
43. Vandierendonck A, Kemps E, Fastame MC, Szmalec A. Working memory components of the Corsi blocks task. Br J Psychol (London England: 1953). 2004;95(Pt 1):57–79.
44. Bowie CR, Harvey PD. Administration and interpretation of the Trail Making Test. Nat Protoc. 2006;1(5):2277–81.
45. Hort J, Laczó J, Vyhnálek M, Bojar M, Bureš J, Vlček K. Spatial navigation deficit in amnestic mild cognitive impairment. *Proceedings of the National Academy of Sciences* 2007, 104(10):4042–4047%U https://www.pnas.org/doi/4010.1073/pnas.0611314104
46. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med. 2016;15(2):155–63.
47. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. J Strength Cond Res. 2005;19(1):231–40.
48. Principles. and practice of structural equation modeling, 4th ed. In: *Principles and practice of structural equation modeling, 4th ed.* New York, NY, US: Guilford Press; 2016: xvii, 534-xvii, 534.
49. Bentler PM. Comparative fit indexes in structural models. Psychol Bull. 1990;107(2):238–46.
50. Hu L-t, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Struct Equ Model. 1999;6(1):1–55.
51. Tucker LR, Lewis C. A reliability coefficient for maximum likelihood factor analysis. Psychometrika. 1973;38(1):1–10.
52. Calamia M, Markon K, Tranel D. The robust reliability of neuropsychological measures: meta-analyses of test-retest correlations. Clin Neuropsychol. 2013;27(7):1077–105.
53. Enkavi AZ, Eisenberg IW, Bissett PG, Mazza GL, MacKinnon DP, Marsch LA, Poldrack RA. Large-scale analysis of test–retest reliabilities of self-regulation measures. *Proceedings of the National Academy of Sciences* 2019, 116(12):5472–5477.
54. Goldberg TE, Harvey PD, Wesnes KA, Snyder PJ, Schneider LS. Practice effects due to serial cognitive assessment: Implications for preclinical Alzheimer's disease randomized controlled trials. Alzheimers Dement (Amst). 2015;1(1):2352–8729.
55. Haith AM, Krakauer JW. The multiple effects of practice: skill, habit and reduced cognitive load. Curr Opin Behav Sci. 2018;20:196–201.
56. Ebel RL. Procedures for the Analysis of Classroom Tests. Educ Psychol Meas. 1954;14(2):352–64.
57. Davies G, Lam M, Harris SE, Trampush JW, Luciano M, Hill WD, Hagenaars SP, Ritchie SJ, Marioni RE, Fawns-Ritchie C, et al. Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. Nat Commun. 2018;9(1):2098.
58. Gonthier C. Cross-cultural differences in visuo-spatial processing and the culture-fairness of visuo-spatial intelligence tests: an integrative review and a model for matrices tasks. Cogn Res Princ Implic. 2022;7(1):11.
59. Salthouse TA. When does age-related cognitive decline begin? Neurobiol Aging. 2009;30(4):507–14.
60. Lovden M, Fratiglioni L, Glymour MM, Lindenberger U, Tucker-Drob EM. Education and Cognitive Functioning Across the Life Span. Psychol Sci Public Interest. 2020;21(1):6–41.
61. LeMoult J, Gotlib IH. Depression: A cognitive perspective. Clin Psychol Rev. 2019;69:51–66.
62. McCloskey G, Hewitt J, Henzel JN, Eusebio E. Executive functions and emotional disturbance. Emotional disorders: A neuropsychological, psychopharmacological, and educational perspective. edn. Middletown, MD, US: School Neuropsych; 2009. pp. 65–105.
63. Dalgleish T, Black M, Johnston D, Bevan A. Transdiagnostic approaches to mental health problems: Current status and future directions. J Consult Clin Psychol. 2020;88(3):179–95.
64. Koo BM, Vizer LM. Mobile Technology for Cognitive Assessment of Older Adults: A Scoping Review. Innov Aging. 2019;3(1):igy038.
65. Willis VC, Thomas Craig KJ, Jabbarpour Y, Scheufele EL, Arriaga YE, Ajinkya M, Rhee KB, Bazemore A. Digital Health Interventions to Enhance Prevention in Primary Care: Scoping Review. JMIR Med Inf. 2022;10(1):e33518.
66. Mulligan CA, Ayoub JL. Remote Assessment: Origins, Benefits, and Concerns. J Intell 2023, 11(6).
67. Leutner F, Codreanu SC, Brink S, Bitsakis T. Game based assessments of cognitive ability in recruitment: Validity, fairness and test-taking experience. Front Psychol. 2022;13:942662.

Publisher's note