HHS Public Access

Author manuscript

Schizophr Res. Author manuscript; available in PMC 2016 December 01.

Published in final edited form as:

Schizophr Res. 2015 December ; 169(1-3): 109–115. doi:10.1016/j.schres.2015.09.003.

The MATRICS Consensus Cognitive Battery (MCCB): Co-norming and standardization in China

Chuan Shi[a,b,c], **Lan Kang**[a,b,c], **Shuqiao Yao**[d], **Yibin Ma**[a,b,c], **Tao Li**[a,b,c], **Ying Liang**[a,b,c], **Zhang Cheng**[a,b,c], **Yifeng Xu**[e], **Jianguo Shi**[f], **Xiufeng Xu**[g], **Congpei Zhang**[h], **Donald R. Franklin**[i], **Robert K. Heaton**[i], **Hua Jin**[i,*], and **Xin Yu**[a,b,c,**]

[a] Peking University Institute of Mental Health (Sixth Hospital), Beijing, China

[b] National Clinical Research Center for Mental Disorders (Peking University Sixth Hospital) Beijing, China

[c] Key Laboratory of Mental Health, Ministry of Health (Peking University), Beijing, China

[d] The Medical Psychological Research Center, Second Xiangya Hospital, Central South University, Changsha, Hunan, China

[e] Division of Psychotic Disorders, Shanghai Mental Health Center, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[f] Xian Mental Health Center, New Qujiang District, Xian, Shanxi, China

[g] Department of Psychiatry, First Affiliated Hospital of Kunming Medical University, Kunming, Yunnan, China

[h] The First Haerbin Psychiatric Hospital, Haerbin, Heilongjiang, China

[i] Department of Psychiatry, University of California at San Diego, San Diego, CA, USA

Abstract

MATRICS Consensus Cognitive Battery (MCCB), packaging 10 tests selected from more than 90 nominated tests, is a method developed by the Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) group to evaluate the efficacy of treatments targeting cognitive impairments in schizophrenia. MCCB had been translated into a number of languages, but only the US and Spain had normative data reported. Inconsistency in translation and cultural differences make direct application of MCCB in China problematic. In this study, we administered the battery to a representative community sample based on Chinese population census in 2005 and obtained normative data. The effects of age, gender, education level, and scale of residence area on

[*] Corresponding author hjin@ucsd.edu (H. Jin). [**] Correspondence to: X. Yu, Peking University Institute of Mental Health (Sixth Hospital), National Clinical Research Center for Mental Disorders (Peking University Sixth Hospital), Key Laboratory of Mental Health, Ministry of Health (Peking University), No.51 Huayuanbei Road, Beijing 100191, China, yuxin@bjmu.edu.cn (X. Yu).

Contributors

SC, YX, YSQ, JH and RKH designed the study and wrote the protocol. SC and KL managed the literature searches and analyses. MYB, LY, CZ, XYF, SJG, ZCP and XXF selected the sample, evaluated patients and contributed in some aspects of the study design and in the interpretation of results. SC and DRF undertook the statistical analysis. SC, YX, and JH wrote the first draft of the manuscript. All authors contributed to and have approved the final manuscript.

Conflict of interest

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

test performance were examined. The sample included 656 healthy volunteers from six sites in China. At each site, sample was stratified according to age, gender, and educational level, and scale of the area one was born in, grew up in and currently living in was recorded. We found age, gender, and education had significant effects on the normative data for MCCB in China, which are comparable to those found for the original standardized English version in the U.S. and the Spanish version in Spain. Remarkably, the residence scale effects on neuropsychological performance were significant, which should be taking into account when calculating the standardized T score for each subject. The practice effects were minor and test-retest reliability of MCCB was good, which suggests MCCB as an appropriate measure for clinical and research usage in China.

Keywords

Schizophrenia; Cognition; MATRICS; MCCB

1. Background

Among the core features of schizophrenia, cognitive impairments occur in prodrome stage, first episode, as well as in chronic stage (Bliksted et al., 2014; Pietrzak et al., 2009; Simon et al., 2012). The cognitive effects concerned generally include traditional cognitive abilities (e.g., executive functions, processing speed, and episodic memory) and social cognition, and those features are crucial predictors of functional outcome in schizophrenia (Fett et al., 2011; Torgalsboen et al., 2014). Mental health professionals therefore have attempted to enhance cognition in schizophrenia through a variety of methods, by administration of typical and atypical antipsychotics, cognition-enhancing agents, cognitive remedy therapy (Wykes et al., 2011), and repetitive transcranial magnetic stimulation (rTMS). However, the results so far were not uniform in the field. Although the meta-analysis by Woodward et al. (Woodward et al., 2005) shows a slight benefit of second generation over first generation medications on cognitive effects, other studies such as CATÉ (Keefe et al., 2007a) in the United States and CAFÉ (Keefe et al., 2007b) in Europe failed to find such a trend. The effects of new cognition-enhancing agents on cognition in schizophrenia, like α7-Nicotinic receptor agonists, Dopamine D1 receptor agonists and AMPA glutamatergic receptor agonists have been supported by preliminary evidence, and studies on this approach are in progress (Buchanan et al., 2007; Marder, 2006, 2011). The mean effect size of cognitive remediation therapy in schizophrenia was reported to be 0.45, but the generalizability of training effects requires further investigation (Wykes et al., 2011). Most of the few studies of rTMS effects on cognition focused on major depressive disorder. Among the rest, one study suggested that rTMS could improve working memory in schizophrenia, whereas another study had a negative outcome (Barr et al., 2013; Guse et al., 2013). A number of studies that aimed to improve cognition but ended up with mixed results might be a result of lack of uniform method to measure cognitive impairments.

More than 130 scientists from academia, government, and the pharmaceutical industry met for the initial MATRICS consensus conference in 2004 to develop a consensus neurocognitive test battery to evaluate the improvement of cognition in schizophrenia.

During the conference, an agreement on the seven cognitive domains for the battery and the five criteria for co-norming test selection has been reached (Green et al., 2004). The finalized MATICS Consensus Cognitive Battery (MCCB) included 10 renowned subtests grouped by cognitive domains (Nuechterlein et al., 2008). Although normative data for these subtests already existed, the MATRICS Neurocognition Committee still recommended a standardization and co-norming process to facilitate interpretation of the results of MCCB (Kern et al., 2004). Accordingly, researchers in the US and Spain obtained normative data in their countries respectively, and they found age, gender and education had significant effects on the normative data (Kern et al., 2008; Rodriguez-Jimenez et al., 2012).

MCCB had been translated into and becoming commercially available in 21 languages, including simplified Chinese, German, Hindi, Russian, Spanish, etc. With the permission of Keith Nuechterlein, the President of MATRICS Assessment, Inc. and Michael Green, the developer of MCCB from University of California, Los Angeles (UCLA), the Chinese MCCB translation team, consisted of Xin Yu and Chuan Shi from Institute of Mental Health, Peking University (IMHPU), Hua Jin and Robert Heaton from University of California, San Diego (UCSD), finished the translation of the MCCB administration and scoring manual and record form in April 2008, then a translation company from Britain conducted the MCCB back translation. In September 2008, Keith Nuechterlein signed the contract with Xin Yu to establish the China MCCB norms. The contract provided authorization and part of funding for the MCCB norm study in China.

The normative data from different countries varied, even for countries that have same official language. For example, a sample of 117 Chinese, Malay and Indian English speakers from Singapore showed poorer performance compared to the US sample, and it may be due to the group differences in cultural and educational backgrounds (Rapisarda et al., 2013). Differences had also been found among African Americans, Japanese and Caucasians with the Cognitive Abilities Screening Instrument (Shadlen et al., 2001). It is reasonable to expect translation and cultural modification to generate even larger differences (Hsieh and Tori, 2007; Shuttleworth-Edwards et al., 2004), and that makes a broader collection of normative data from different cultures necessary for a wider use of MCCB. Therefore, we collected MCCB normative data in Mainland China.

2. Materials and methods

2.1. Participants

The sample was recruited in six different sites in China by six institutions: Beijing (the capital of China) by IMHPU; Changsha (Middle East of China) by the Second Xiangya Hospital, Central South University; Shanghai (South East of China) by Shanghai Mental Health Center, Shanghai Jiao Tong University; Xi'an (North West of China) by Xi'an Mental Health Center; Harbin (North East of China) by the First Harbin Psychiatric Hospital; and Kunming (South West of China) by the First Affiliated Hospital of Kunming Medical University. The six sites were chosen to represent six geographic regions of China, and the sample was stratified by age, gender and education as reported by China Census Bureau in 2005 (China Census Bureau, 2006). Four age groups: 20–29, 30–39, 40–49 and 50–59, were considered by the corresponding proportions (25.2%, 25.4%, 25.1% and 24.3%) of subjects according to the age distribution of Chinese population. 50.3% of the subjects were male and 49.7% were female. Following the Chinese educational system, the subjects were grouped as having completed primary studies (56.2%), secondary studies (23.1%), and university studies (20.7%). A total of 656 healthy volunteers were assessed with MCCB for China. After 4–8 weeks, we retested 186 subjects, 93 males and 93 females. The participants received 50 RMB for transportation each time they attended to the study. We used same forms of NAB Mazes, HVLT-R and BVMT-R in the retest session.

2.2. Recruitment procedures

The recruitment procedure was identical in all six sites. In each site, the potential subjects were informed of the study through referrals and information sheets in the community, and those interested contacted the recruiters to schedule an in-person interview, in which the subjects were screened with the inclusion and exclusion criteria. The inclusion criteria were: 20–59 years old; at least 5 years of education; capable of reading and understanding Mandarin. If potential subjects fit the inclusion criteria, the recruiters explained the study and gave the informed consent to them and subjects signed the document if agreed to participate, and then were consecutively put into the respective subgroups according to their age, gender, and education level until the recruitment cells were filled. The participants did not get reimbursement other than 50 RMB for transportation. All procedures were approved by the IRB of IMHPU.

Exclusion criteria were similar to the original MCCB study (Kern et al., 2008): history of diagnosis of schizophrenia or other psychotic disorder; history of other major psychiatric disorders; diagnosis of mental retardation or pervasive developmental disorder; clinically significant neurological disease or head injury with loss of consciousness; currently taking psychiatric drugs, cognition-enhancing medication, narcotics for pain or other medications which could potentially affect cognitive performance; history of alcohol or drug abuse/ dependence; any substance use in the last 3 days prior to testing; more than four alcoholic drinks per day for each of the last 3 days prior to testing; inability to understand Mandarin sufficiently to comprehend testing instructions; and inability to comprehend the consent form appropriately.

2.3. Study design

Certified neuropsychological examiners in each site conducted a structured interview to acquire the subjects' demographic data then administered the MCCB tests in following order:

1. Trail Making Test: Part A (TMT)
2. Brief Assessment of Cognition in Schizophrenia (BACS): Symbol Coding
3. Hopkins Verbal Learning Test—Revised (HVLT-R)
4. Wechsler Memory Scale—Third Edition (WMS-III): Spatial Span
5. Neuropsychological Assessment Battery (NAB): Mazes
6. Brief Visuospatial Memory Test—Revised (BVMT-R)

7. Category Fluency Test: Animal naming (Fluency)
8. Mayer–Salovey–Caruso Emotional Intelligence Test (MSCEIT): Managing Emotions
9. Continuous Performance Test—Identical Pairs (CPT-IP).

The HVLT-R and BVMT-R were only administered for three learning trials. The Letter-Number Span Test (LNS) was excluded from original MCCB because there is no corresponding alphabet in Chinese. We also made some cultural modification in the translation of HVLT-R, for example: coin substituted penny, because Chinese population are more familiar with coin; Er Hu substituted violin, because Er Hu is a string instrument equally known in urban and rural areas in China, but mostly only urban population are familiar with violin; He Shang (Buddhist Monk) substituted priest, because Chinese population are generally familiar with monks but priests are only common in certain regions. We also changed the Western names into Chinese names in the MSCEIT™ Managing Emotions (Branch 4), such as: Liu Jia for Mara, Hu Wei for Andrew, and Zhou Hua for Jane.

2.4. Data analysis

Normality of distribution was first examined for each of the nine MCCB tests, and those that were notably skewed were transformed. We used a logarithmic transformation for data skewed to the left of the Gaussian curve, and a $y = x^2$ transformation for data skewed to the right. In reference to Kern et al. (2008), the raw scores or transformed raw scores for the 9 tests were standardized to scaled scores based on 656 Chinese participants so that each test had a mean of 10 and a SD of 3 in the sample. The composite scaled score was generated from taking the average score of 9 subtests scaled scores and re-scaled so that it also had a mean of 10 and a SD of 3.

The scaled scores for the seven cognitive domains and the composite scaled score were analyzed to examine age, gender, education and residence scale effects. The residence scale effects were tested by the area one was born in, grew up in, and currently living in. The residence areas were coded according to same criteria, 1 for rural area, 2 for small city (population less than 1 million), 3 for middle city (population 1 to 5 million), and 4 for large city (population more than 5 million). Differences in performance related to age, education and residence scale were assessed with ANCOVA, whereas independent t-test was used to evaluate gender effect. The effect size (Cohen's d) was calculated for practice effects and Intraclass Correlation Coefficient (ICC) for test-retest reliability.

3. Results

3.1. Demographic characteristics

The average age of the participants was 39.3 years, and the average education level was 10.8 years, which is in the middle of secondary school, and the gender ratio was about half to half. Among the subjects, 77.4% of them were married, 19.1% were never married, 2.6% were divorced, and 0.9% were widows. Consistent with the migration patterns in China, the variation between the area one was born in and grew up in was not very large in the sample, however, where one grew up and where one currently living varied greatly: few people chose to stay in rural areas, and the proportion of rural residents has decreased from 26.8% to 5.2%. In turn, many people moved to big cities, the proportion increased from 45.3% to 61.6% (see Table 1).

3.2. Age effects

Since the education level differs among four age groups, education was taken as a covariate when conducting ANCOVAs for age effects. A significant age effect was found for all nine tests, as well as the Composite Scaled Score (Trail making A: F = 52.61, df = 3,652, P < 0.001; BACS Symbol Coding: F = 125.32, df = 3,652, P < 0.001; HVLT-R learning: F = 47.48, df = 3,652, P < 0.001; WMS-III Spatial Span: F = 24.77, df = 3,652, P < 0.001; NAB Mazes: F = 14.99, df = 3,652, P < 0.001; BVMT-R Learning: F = 48.36, df = 3,652, P < 0.001; Animal Fluency: F = 14.52, df = 3,652, P < 0.001; MSCEIT Managing Emotions: F = 11.20, df = 3,652, P < 0.001; CPT-IP: F = 44.29, df = 3,652, P < 0.001; Composite Scaled Score: F = 146.10, df = 3,652, P < 0.001) (see Table 2).

The pattern of age effects was similar for all nine cognitive tests: young participants consistently performed better than the older, and the effects persisted when education level was controlled. Bonferroni-corrected post-hoc analyses revealed significant differences across the four age groups for Trail making A, BACS Symbol Coding, HVLT-R, NAB Mazes, BVMT-R and Overall Composite Scaled Score, with best performance in the 20–29 year-old group, good performance in the 30– 39 year-old group, poorer performance in the 40–49 year-old group, and the poorest performance in the 50–59 year-old group. In the Animal Fluency subtest, the group aged 20–29 and 30–39 performed significantly better than the 40–49 and 50–59 groups, with no significant differences between the latter two. We observed a similar pattern for the subtests of WMS-III Spatial Span and MSCEIT Managing Emotions. In the CPT-IP subtest, the groups aged 20–29 and 30–39 both performed significantly better than the 40–49 and 50–59 groups, the group aged 40–49 performed significantly better than 50–59 group, with no significant difference between 20–29 and 30–29 group.

3.3. Gender effects

Several cognitive subtests showed gender effects. In the WMS-III Spatial Span and NAB Mazes tests, men performed better than women (t = 2.13, df = 654, p = 0.034, cohen's d = 0.17; t = 4.28,df = 654, p < 0.001, cohen's d = 0.33), whereas women performed better than men on BACS Symbol Coding and MSCEIT Managing Emotions (t = 3.50, df = 654, p < 0.001, cohen's d = 0.27; t = 2.08, df = 654, p = 0.038, cohen's d = 0.16). No statistically significant difference was found between men and women on the rest of cognitive tests, or on the Overall Composite Scaled Score (see Table 3).

3.4. Education effects

Since the age distributions were different among three education level groups, age covaried when conducting ANCOVAs to examine education effects. Education had a significant effect on all nine tests and the Composite Scaled Score (Trail making A: F = 17.71, df = 2,653, P < 0.001; BACS Symbol Coding: F = 46.11, df = 2,653, P < 0.001; HVLT-R Learning: F = 47.05, df = 2,653, P < 0.001; WMS Spatial Span: F = 10.13, df = 2,653, P < 0.001; NAB

Mazes: $F = 20.35$, $df = 2,653$, $P < 0.001$; BVMT-R Learning: $F = 28.91$, $df = 2,653$, $P < 0.001$; Animal Fluency: $F = 18.28$, $df = 2,653$, $P < 0.001$; MSCEIT Managing Emotions: $F = 10.80$, $df = 2,653$, $P < 0.001$; CPT-IP: $F = 25.54$, $df = 2,653$, $P < 0.001$; Composite Scaled Score: $F = 193.46$, $df = 2,653$, $P < 0.001$) (see Table 4).

Across the nine cognitive subtests, subjects with more than 12 years of education performed the best, those with 9–12 years education had an intermediate performance, and those with only preliminary school education performed poorest. Bonferroni-corrected post-hoc analyses revealed significant differences among all three education levels for the all nine cognitive subtests and Overall Composite Scaled Score.

3.5. Residence scale effects

When we used composite scaled score to explore residence scale effects, we found the area one was born in ($F = 110.1$, $df = 3,652$, $p < 0.001$), the area one grew up in ($F = 113.8$, $df = 3,652$, $p < 0.001$) and the area one current living in ($F = 104.0$, $df = 3,652$, $p < 0.001$) all had significant effects even as controlling for age, gender and education. Bonferroni-corrected post-hoc analyses revealed participants that grew up in cities performed better than those grew up in rural areas, and participants that currently living in big cities performed better than those living in small cities and rural areas, participants that currently living in middle-size cities also performed better than those in rural areas.

3.6. Practice effects and test-retest reliability

Substantial practice effects were noted in MCCB. However, they varied considerably across the nine subtests, with effect sizes ranging from minimal and not statistically significant (Animal fluency and MCSEIT Managing Emotion) to large (HVLT-R and BVMT-R Learning). In terms of test-retest reliability, the ICC of nine cognitive subtests varied from 0.73 to 0.94, the lowest being MCSEIT Managing Emotion (ICC = 0.73) and the highest being BACS Symbol Coding (ICC = 0.94). The ICC for whole test battery (Composite score) was 0.95 (see Table 5).

4. Discussion

Consistent with the original standardized English version of MCCB in the US and the Spanish version in Spain, we found age, gender and education had significant influence on the normative data in China, and remarkably, so did scale of residence area. Our study also demonstrates that all components of MCCB have good reliability in China. However, the majority of MCCB subtests are subject to substantial practice effects, and that need to be considered when measuring disease- or treatment-driven cognitive effects changes with the battery in the long term.

With respect to effects of age and education, which were evident for all the seven subdomains in MCCB, older in age and lower level of education were associated with worse cognitive performance. Although no significant differences across 20–39, 40–49 and 50–59 years groups for the verbal learning test were observed in the US (Kern et al., 2008), our results are similar to normative data in Spain (Rodriguez-Jimenez et al., 2012), where the 20–29 and 30–39 years groups had better verbal memory compared to the 40–49 and 50–59 years groups. We also found significant differences between the 20–29 and 30–39 years groups in multiple cognitive subdomains such as speed of information processing, verbal and visual learning, and executive function, and therefore it is reasonable to divide the age group from 20 to 59 into four. In contrast to the US normative data, younger participants in China performed better than the older, whereas older participants in the US performed better than the younger in social cognition domain (Kern et al., 2008). The content of MSCEIT Branch 4 might help with understanding the difference, for example, some items related to driving and changing job are more relevant to younger than older people in China. Comparing the Chinese and the Spanish normative data, similar pattern in social cognition domain was found (Rodriguez-Jimenez et al., 2012). Social cognition appeared to be the most culturally “biased” (affected) part of MCCB, as a result now there is an international scoring for the MSCEIT Branch 4 for MCCB that eliminated some items that were inconsistent across languages and cultures. The most consistent finding across countries is the substantial education effects, for which three countries had the same outcomes (Kern et al., 2008; Rodriguez-Jimenez et al., 2012).

As for gender, male subjects getting higher scores in reasoning, problem solving and working memory in our study were similar to the results from the original US normative study and Spanish normative study (Kern et al., 2008; Rodriguez-Jimenez et al., 2012). Another subdomain performed better by males in Spain was attention/vigilance, which was not found in our study and US normative study (Kern et al., 2008; Rodriguez-Jimenez et al., 2012). Other studies on typical population and schizophrenia patients showed that gender had no significant effect on CPT performance (Kao et al., 2013; Markovska-Simoska and Pop-Jordanova, 2009). In the US normative study, women performed better than men in verbal learning, whereas in our study there was only a non-significant trend of the same direction ($p = 0.065$, Cohen's $d = 0.15$) (Kern et al., 2008). The other two subdomains where women stood out were speed of information processing and social cognition in our study, similar finding was only observed in social cognition subdomain in Spain, while in the US there was only a non-significant tendency. Neither the US nor Spain normative study exhibited gender difference in speed of information processing (Kern et al., 2008; Rodriguez-Jimenez et al., 2012). In general, females had higher mental processing speed and verbal ability, males had better visuospatial ability, which were demonstrated in both China and the US (Halpern, 1997; Harrison and Whissell, 1980), and females in China and Spain showed better emotion perception and regulation but not in the US.

Residence scale effects were robust in our study, participants grew up and currently living in rural areas had poorer neurocognition performance compared to those in cities. The different performance of rural and urban residents was also observed in our previous study, which is probably a result of difference in income level, education opportunity and quality, professional and academic practice opportunity, and diversity of environmental stimuli, etc. (Gupta et al., 2011; Tine, 2014). More studies are in need to further infer the causes.

Practice effects are big concerns for neuropsychological tests, particularly for clinical trials that require multiple visits, thus the MATRICS group set less practice effects as a criterion when selecting MCCB candidate tests (Nuechterlein et al., 2008). Except for the two learning tests, the rest of MCCB subtests had small to medium practice effects in China, although they were not as good as in the study by Nuechterlein et al. (2008) in US (Cohen's d from 0.02 to 0.49 in China vs. 0.00 to 0.22 in US). The difference could be attributed to performance of the healthy sample used in our study would improve faster through practice than the patient sample in Nuechterlein et al. (2008) study. In regard to the two learning tests, using the same form twice in China caused high practice effects, which means alternate form for verbal and visual learning tests in the follow-up visits are crucial for controlling the practice effects. For the NAB Mazes, although the practice effects were medium (Cohen's d = 0.43), we also recommend alternate form in multiple visits study. For test-retest reliability, although our test interval was longer than US study (4–8 weeks vs. 4 weeks), the reliability shown in our study was better than that in the US study (ICC from 0.73 to 0.94 in China vs. 0.68 to 0.85 in US) (Nuechterlein et al., 2008), which might also be a result of healthy and patient sample differences.

Usually the standardized T score for neuropsychological tests is generated adjusting by age, gender and education (Hsieh and Tori, 2007; Kern et al., 2008; Rodriguez-Jimenez et al., 2012), whereas in our study, we found that scale of the areas where one grew up and currently living also mattered in neuropsychological performance. Thus, establishing the demographically adjusted equations for T scores in China also needs to consider the residence scale effects in order to yield the precise scores. Future research needs to investigate whether the effects apply to other countries. In China, cities provide more learning opportunity for children than rural areas do, and children in cities have a greater variety of extracurricular activities. Therefore, it was not surprising that even with the same education level, people grew up in cities performed better on neuropsychological tests than those in rural areas. Environmental enrichment also was observed to promote adult brain plasticity (Sale et al., 2014). On the other hand, urban environment could also be a source of stress: in a study by Lederbogen et al. (2011), adults grew up in cities showed more dysregulated frontal cognitive control while performing stressful mental arithmetic than those grew up in rural areas. More cross-cultural studies are in need to interpret this finding.

One limitation of the present study is that we did not have a substitute subtest for Letter-Number Sequencing, and therefore the test to evaluate verbal working memory in Chinese MCCB is missing; digit span or the paced auditory serial addition task could be considered as supplemental tests covering this domain. Another limitation is that we did not test the norms acquired from the US and Spain in China, but judging from the few previously stated results differences among our study and theirs, we propose never automatically assume equivalence of norms cross cultures. With the cross-sectional norms obtained in this study, future studies should target at establishing longitudinal norms (norms for change).

In summary, the present study presents norms of MCCB obtained from native speakers of Mandarin, the most widely used language in the world. Our sample represented the Chinese population well and was large enough to be compared with the results of US and Spain norms (Nuechterlein et al., 2008; Rodriguez-Jimenez et al., 2012).The psychometric properties of Chinese MCCB are good, in reference to those reported from the US and Spain, which supported the potential application of MCCB in different countries. The present co-norming and standardization study for MCCB in a nation-wide Chinese sample enables the clinical and research usage of this valuable cognitive test battery in China.

Acknowledgments

We thank all the staff in six study sites who involved in organizing and recruiting samples as well as all the subjects providing the neurobehavioral data to our norms.

Role of funding resource

This research was supported in part by the National Key Project of Scientific and Technical Supporting Programs Funded by Ministry of Science & Technology of China (No. 2007BAI17B04).

References

Barr MS, Farzan F, Rajji TK, Voineskos AN, Blumberger DM, Arenovich T, Fitzgerald PB, Daskalakis ZJ. Can repetitive magnetic stimulation improve cognition in schizophrenia? Pilot data from a randomized controlled trial. Biol. Psychiatry. 2013; 73(6):510–517. [PubMed: 23039931]

Bliksted V, Fagerlund B, Weed E, Frith C, Videbech P. Social cognition and neurocognitive deficits in first-episode schizophrenia. Schizophr. Res. 2014; 153(1–3):9–17. [PubMed: 24480016]

Buchanan RW, Freedman R, Javitt DC, Abi-Dargham A, Lieberman JA. Recent advances in the development of novel pharmacological agents for the treatment of cognitive impairments in schizophrenia. Schizophr. Bull. 2007; 33(5):1120–1130. [PubMed: 17641146]

China Census Bureau, C. China Population in 2005. China Statistics Press; Beijing: 2006.

Fett AKJ, Viechtbauer W, Dominguez MD, Penn DL, van Os J, Krabbendam L. The relationship between neurocognition and social cognition with functional outcomes in schizophrenia: a meta-analysis. Neurosci. Biobehav. Rev. 2011; 35(3):573–588. [PubMed: 20620163]

Green MF, Nuechterlein KH, Gold JM, Barch DM, Cohen J, Essock S, Fenton WS, Frese F, Goldberg TE, Heaton RK, Keefe RS, Kern RS, Kraemer H, Stover E, Weinberger DR, Zalcman S, Marder SR. Approaching a consensus cognitive battery for clinical trials in schizophrenia: the NIMH-MATRICS conference to select cognitive domains and test criteria. Biol. Psychiatry. 2004; 56(5):301–307. [PubMed: 15336511]

Gupta S, Vaida F, Riggs K, Jin H, Grant I, Cysique L, Shi C, Yu X, Wu Z, Heaton RK, Group, H.I.V.N.R.C. Neuropsychological performance in mainland China: the effect of urban/rural residence and self-reported daily academic skill use. J. Int. Neuropsychol. Soc.: JINS 17. 2011; 1:163–173.

Guse B, Falkai P, Gruber O, Whalley H, Gibson L, Hasan A, Obst K, Dechent P, McIntosh A, Suchan B, Wobrock T. The effect of long-term high frequency repetitive transcranial magnetic stimulation on working memory in schizophrenia and healthy controls–a randomized placebo-controlled, double-blind fMRI study. Behav. Brain Res. 2013; 237:300–307. [PubMed: 23022750]

Halpern DF. Sex differences in intelligence. Implications for education. Am. Psychol. 1997; 52(10): 1091–1102. [PubMed: 9329293]

Harrison B, Whissell CM. Neuroticism, practice, and sex differences in a digit-symbol task. Percept. Mot. Skills. 1980; 50(2):487–490. [PubMed: 7375300]

Hsieh SL, Tori CD. Normative data on cross-cultural neuropsychological tests obtained from Mandarin-speaking adults across the life span. Arch. Clin. Neuropsychol.: The Official Journal of the National Academy of Neuropsychologists. 2007; 22(3):283–296.

Kao YC, Liu YP, Lien YJ, Lin SJ, Lu CW, Wang TS, Loh CH. The influence of sex on cognitive insight and neurocognitive functioning in schizophrenia. Prog. Neuro-Psychopharmacol. Biol. Psychiatry. 2013; 44:193–200.

Keefe RS, Bilder RM, Davis SM, Harvey PD, Palmer BW, Gold JM, Meltzer HY, Green MF, Capuano G, Stroup TS, McEvoy JP, Swartz MS, Rosenheck RA, Perkins DO, Davis CE, Hsiao JK, Lieberman JA, Investigators C. Arch. Gen. Psychiatry. 64(6):633–647. [PubMed: 17548746]

Keefe RS, Sweeney JA, Gu H, Hamer RM, Perkins DO, McEvoy JP, Lieberman JA. Effects of olanzapine, quetiapine, and risperidone on neurocognitive function in early psychosis: a randomized, double-blind 52-week comparison. Am. J. Psychiatry. 2007b; 164(7):1061–1071. [PubMed: 17606658]

Kern RS, Green MF, Nuechterlein KH, Deng BH. NIMH-MATRICS survey on assessment of neurocognition in schizophrenia. Schizophr. Res. 2004; 72(1):11–19. [PubMed: 15531403]

Kern RS, Nuechterlein KH, Green MF, Baade LE, Fenton WS, Gold JM, Keefe RS, Mesholam-Gately R, Mintz J, Seidman LJ, Stover E, Marder SR. The MATRICS Consensus Cognitive Battery, part 2: co-norming and standardization. Am. J. Psychiatry. 2008; 165(2):214–220. [PubMed: 18172018]

Lederbogen F, Kirsch P, Haddad L, Streit F, Tost H, Schuch P, Wust S, Pruessner JC, Rietschel M, Deuschle M, Meyer-Lindenberg A. City living and urban upbringing affect neural social stress processing in humans. Nature. 2011; 474(7352):498–501. [PubMed: 21697947]

Marder SR. The NIMH-MATRICS project for developing cognition-enhancing agents for schizophrenia. Dialogues Clin. Neurosci. 2006; 8(1):109–113. [PubMed: 16640121]

Marder SR. Lessons from MATRICS. Schizophr. Bull. 2011; 37(2):233–234. [PubMed: 21325472]

Markovska-Simoska S, Pop-Jordanova N. Comparison of visual and emotional continuous performance test related to sequence of presentation, gender and age. Prilozi/ Makedonska akademija na naukite i umetnostite, Oddelenie za bioloski i medicinski nauki = Contributions/ Macedonian Academy of Sciences and Arts. Sect. Biol. Med. Sci. 2009; 30(1):167–178.

Nuechterlein KH, Green MF, Kern RS, Baade LE, Barch DM, Cohen JD, Essock S, Fenton WS, Frese FJ 3rd, Gold JM, Goldberg T, Heaton RK, Keefe RS, Kraemer H, Mesholam-Gately R, Seidman LJ, Stover E, Weinberger DR, Young AS, Zalcman S, Marder SR. The MATRICS Consensus Cognitive Battery, part 1: test selection, reliability, and validity. Am. J. Psychiatry. 2008; 165(2): 203–213. [PubMed: 18172019]

Pietrzak RH, Snyder PJ, Jackson CE, Olver J, Norman T, Piskulic D, Maruff P. Stability of cognitive impairment in chronic schizophrenia over brief and intermediate re-test intervals. Hum. Psychopharmacol. 2009; 24(2):113–121. [PubMed: 19090506]

Rapisarda A, Lim TF, Lim M, Collinson SL, Kraus MS, Keefe RS. Applicability of the MATRICS Consensus Cognitive Battery in Singapore. Clin. Neuropsychol. 2013; 27(3):455–469. [PubMed: 23362823]

Rodriguez-Jimenez R, Bagney A, Garcia-Navarro C, Aparicio AI, Lopez-Anton R, Moreno-Ortega M, Jimenez-Arriero MA, Santos JL, Lobo A, Kern RS, Green MF, Nuechterlein KH, Palomo T. The MATRICS consensus cognitive battery (MCCB): co-norming and standardization in Spain. Schizophr. Res. 2012; 134(2–3):279–284. [PubMed: 22192501]

Sale A, Berardi N, Maffei L. Environment and brain plasticity: towards an endogenous pharmacotherapy. Physiol. Rev. 2014; 94(1):189–234. [PubMed: 24382886]

Shadlen MF, Larson EB, Gibbons LE, Rice MM, McCormick WC, Bowen J, McCurry SM, Graves AB. Ethnicity and cognitive performance among older African Americans, Japanese Americans, and Caucasians: the role of education. J. Am. Geriatr. Soc. 2001; 49(10):1371–1378. [PubMed: 11890499]

Shuttleworth-Edwards AB, Kemp RD, Rust AL, Muirhead JG, Hartman NP, Radloff SE. Cross-cultural effects on IQ test performance: a review and preliminary normative indications on WAIS-III test performance. J. Clin. Exp. Neuropsychol. 2004; 26(7):903–920. [PubMed: 15742541]

Simon AE, Gradel M, Cattapan-Ludewig K, Gruber K, Ballinari P, Roth B, Umbricht D. Cognitive functioning in at-risk mental states for psychosis and 2-year clinical outcome. Schizophr. Res. 2012; 142(1–3):108–115. [PubMed: 23025995]

Tine M. Working memory differences between children living in rural and urban poverty. J. Cogn. Dev. 2014; 15(4):599–613. [PubMed: 25554726]

Torgalsboen AK, Mohn C, Rishovd Rund B. Neurocognitive predictors of remission of symptoms and social and role functioning in the early course of first-episode schizophrenia. Psychiatry Res. 2014; 216(1):1–5. [PubMed: 24530159]

Woodward ND, Purdon SE, Meltzer HY, Zald DH. A meta-analysis of neuropsychological change to clozapine, olanzapine, quetiapine, and risperidone in schizophrenia. Int. J. Neuropsychopharmacol./Official Scientific Journal of the Collegium Internationale Neuropsychopharmacologicum. 2005; 8(3):457–472.

Wykes T, Huddy V, Cellard C, McGurk SR, Czobor P. A meta-analysis of cognitive remediation for schizophrenia: methodology and effect sizes. Am. J. Psychiatry. 2011; 168(5):472–485. [PubMed: 21406461]

Table 1

Demographic characteristics of norm sample.

	All subjects (N = 656)
Age	39.30 (SD: 11.37)
Education	10.80 (SD:3.18)
Gender	Male 330 (50.3%); female 326 (49.7%)
Marriage	Married 508 (77.4%)
	Widows 6 (0.9%)
	Divorced 17 (2.6%)
	Unmarried 125 (19.1%)
Area the subject was born	Rural area 189 (28.8%)
	Small city 95 (14.5%)
	Middle city 87 (13.3%)
	Big city 285 (43.4%)
Area the subject grew up	Rural area 176 (26.8%)
	Small city 97 (14.8%)
	Middle city 86 (13.1%)
	Big city 297 (45.3%)
Area the subject is current living	Rural area 34 (5.2%)
	Small city 70 (10.7%)
	Middle city 148 (22.6%)
	Big city 404 (61.6%)

Table 2

Mean scaled scores for the 9 MCCB tests, by age group (ANCOVA, education as covariance).

	20–29 years (n = 164)	30–39 years (n = 167)	40–49 years (n = 165)	50–59 years (n = 160)	F/x^2	P
Education	12.01 (SD: 3.03)	10.97 (SD: 3.60)	10.53 (SD: 3.83)	9.66 (SD: 2.73)	16.48	<0.001
Gender						
Male	86	83	82	79	0.40	0.94
Female	78	84	83	81		
Trail Making A	11.96 (SD: 2.64)	10.70 (SD: 2.70)	9.13 (SD: 2.68)	8.13 (SD: 2.56)	52.61	<0.001
BACS Symbol Coding	12.41 (SD: 2.29)	11.10 (SD: 2.44)	8.90 (SD: 2.37)	7.42 (SD: 2.26)	125.32	<0.001
HVLT-R	11.79 (SD: 2.63)	10.65 (SD: 2.78)	9.13 (SD: 2.63)	7.98 (SD: 2.61)	47.48	<0.001
WMS-III Spatial Span	11.50 (SD: 2.65)	10.40 (SD: 2.73)	9.35 (SD: 3.09)	8.65 (SD: 2.57)	24.77	<0.001
NAB Mazes	11.98 (SD: 2.61)	10.52 (SD: 2.78)	9.30 (SD: 2.75)	8.07 (SD: 2.42)	14.99	<0.001
BVMT-R	11.87 (SD: 2.45)	10.86 (SD: 2.92)	9.11 (SD: 2.80)	8.16 (SD: 2.55)	48.36	<0.001
Animal Fluency	11.02 (SD: 2.86)	10.83 (SD: 3.28)	9.53 (SD: 2.62)	8.84 (SD: 2.75)	14.52	<0.001
MSCEIT Managing Emotions	11.11 (SD: 2.79)	10.49 (SD: 2.91)	9.32 (SD: 2.81)	9.31 (SD: 2.88)	11.20	<0.001
CPT-IP	11.56 (SD: 2.55)	10.93 (SD: 2.72)	9.14 (SD: 2.83)	8.17 (SD: 2.46)	44.29	<0.001
Composite Score	12.39 (SD: 2.12)	11.02 (SD: 2.58)	8.89 (SD: 2.47)	7.63 (SD: 2.31)	146.10	<0.001

SD: standard deviation.

Table 3

Mean scaled scores for the 9 MCCB tests, by gender (t-test).

	Male (n = 326)	**Female (n = 330)**	**t**	**P**	**Effect size (Cohen's d)**
Age	39.56 (SD: 11.20)	39.05 (SD: 11.54)	0.57	0.568	0.04
Education	10.79 (SD: 3.22)	10.81 (SD: 3.13)	−0.05	0.963	−0.01
Trail Making A	9.93 (SD: 3.01)	10.05 (SD: 3.03)	0.50	0.614	−0.04
BACS Symbol Coding	9.57 (SD: 2.79)	10.39 (SD: 3.20)	3.50	<0.001	−0.27
HVLT-R	9.68 (SD: 2.90)	10.12 (SD: 3.13)	1.85	0.065	−0.15
WMS-III Spatial Span	10.23 (SD: 2.89)	9.74 (SD: 3.02)	2.13	0.034	0.17
NAB Mazes	10.47 (SD: 2.95)	9.48 (SD: 2.98)	4.28	<0.001	0.33
BVMT-R	9.92 (SD: 2.96)	10.12 (SD: 3.13)	0.85	0.398	−0.07
Animal Fluency	10.17 (SD: 3.09)	9.96 (SD: 2.95)	0.90	0.368	0.07
MSCEIT Managing Emotions	9.82 (SD: 2.77)	10.30 (SD: 3.10)	2.08	0.038	−0.16
CPT-IP	9.92 (SD: 2.96)	10.01 (SD: 2.97)	0.37	0.713	−0.13
Composite Score	9.96 (SD: 2.91)	10.04 (SD: 3.10)	0.32	0.753	−0.03

SD: standard deviation.

Table 4

Mean scaled scores for the 9 MCCB tests, by educational level (ANCOVA, age as covariance).

	<9 years (n = 369)	**9–12 years (n = 151)**	**>12 (n = 136)**	**F/x^2**	**P**
Age	41.76 (SD: 11.00)	38.07 (SD: 11.19)	34.02 (SD: 10.56)	26.01	<0.001
Gender				0.55	0.76
Male	189	72	69		
Female	180	79	67		
Trail Making A	9.29 (SD: 3.04)	10.17 (SD: 2.64)	11.69 (SD: 2.63)	17.71	<0.001
BACS Symbol Coding	9.01 (SD: 2.93)	10.30 (SD: 2.57)	12.25 (SD: 2.40)	46.11	<0.001
HVLT-R	8.91 (SD: 2.79)	10.31 (SD: 2.84)	12.13 (SD: 2.50)	47.05	<0.001
WMS-III Spatial Span	9.39 (SD: 2.87)	10.31 (SD: 2.77)	11.22 (SD: 3.00)	10.13	<0.001
NAB Mazes	9.28 (SD: 2.95)	10.09 (SD: 2.77)	11.77 (SD: 2.67)	20.35	<0.001
BVMT-R	9.14 (SD: 3.02)	10.47 (SD: 2.70)	11.90 (SD: 2.49)	28.91	<0.001
Animal Fluency	9.38 (SD: 3.09)	10.39 (SD: 2.67)	11.56 (SD: 2.61)	18.28	<0.001
MSCEIT Managing Emotions	9.55 (SD: 2.82)	10.23 (SD: 2.88)	11.27 (SD: 3.00)	10.80	<0.001
CPT-IP	9.28 (SD: 2.92)	9.95 (SD: 2.66)	11.86 (SD: 2.60)	25.54	<0.001
Composite Score	8.96 (SD: 2.85)	10.36 (SD: 2.61)	12.50 (SD: 2.19)	193.46	<0.001

SD: standard deviation.

Table 5

Practice effects and test-retest reliability of 9 MCCB tests (n = 185).

	T1	T2	T2–T1 difference	t	P	Effect size (Cohen's d)	ICC
Trail Making A	10.11 (SD: 2.61)	11.46 (SD: 2.89)	1.34 (SD: 2.41)	−7.62	<0.001	−0.49	0.76
BACS Symbol Coding	10.19 (SD: 2.94)	10.68 (SD: 2.87)	0.49 (SD: 1.36)	−4.87	<0.001	−0.17	0.94
HVLT-R	10.13 (SD: 2.82)	12.62 (SD: 3.03)	2.49 (SD: 2.10)	−16.20	<0.001	−0.85	0.85
WMS III Spatial Span	10.07 (SD: 2.83)	10.71 (SD: 3.14)	0.64 (SD: 2.61)	−3.35	0.001	−0.21	0.77
NAB Mazes	10.21 (SD: 2.72)	11.47 (SD: 3.07)	1.26 (SD: 2.18)	−7.84	<0.001	−0.43	0.84
BVMT-R	10.54 (SD: 3.02)	13.32 (SD: 3.54)	2.77 (SD: 2.43)	−15.58	<0.001	−0.84	0.84
Animal Fluency	10.67 (SD: 2.80)	10.83 (SD: 2.81)	0.16 (SD: 2.36)	−0.90	0.37	−0.06	0.78
MSCEIT Managing Emotions	10.19 (SD: 3.15)	10.11 (SD: 3.46)	−0.07 (SD: 3.06)	0.34	0.737	0.02	0.73
CPT-IP	10.72 (SD: 2.97)	11.58 (SD: 2.95)	0.86 (SD: 1.89)	−6.18	<0.001	−0.29	0.89
Composite Score	10.00 (SD: 3.00)	11.69 (SD: 3.20)	1.69 (SD: 1.33)	−17.21	<0.001	−0.55	0.95

SD: standard deviation; ICC: