Published in final edited form as:

J Clin Exp Neuropsychol. 2019 April ; 41(3): 246–256. doi:10.1080/13803395.2018.1529229.

Assessment of Neuropsychological Performance in Mexico City Youth using The Cambridge Neuropsychological Test Automated Battery (CANTAB)

Rivka Green⁽¹⁾, Christine Till^{(1)*}, Hiba Al-Hakeem⁽¹⁾, Rob Cribbie⁽¹⁾, Martha M. Téllez-Rojo⁽²⁾, Erika Osorio⁽³⁾, Howard Hu^{(4),(5)}, Lourdes Schnaas⁽³⁾

⁽¹⁾Department of Psychology, York University, Toronto, Canada

⁽²⁾Centro de Investigación en Nutrición y Salud, Instituto Nacional de Salud Pública, Mexico City, Mexico

⁽³⁾Subdirección de Investigación en Intervenciones Comunitarias, Instituto Nacional de Perinatología, Mexico City, Mexico

⁽⁴⁾Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

⁽⁵⁾University of Washington, USA

Abstract

Background: The Cambridge Neuropsychological Test Automated Battery (CANTAB) is a computerized tool used to measure cognitive function in diverse populations, and is sensitive for assessing developmental changes in children. Although CANTAB has been used in several countries, its applicability in a Mexican child population is unknown. This study examined developmental trends on CANTAB in a large sample of urban Mexico City youth, and tested the hypothesis that their performance would be similar to a large American normative sample.

Methods: As part of a birth cohort, *Early Life Exposures in Mexico to Environmental Toxicants*, 826 children, ages 5–15, completed CANTAB. Subtests measured planning (Stockings of Cambridge; SOC), short-term memory (Delayed Matching to Sample; DMS), sustained attention (Rapid Visual Information Processing; RVP), ability to match visual stimuli (Match to Sample Visual Search; MTS), flexibility (Intra-extra Dimensional Set Shift; IED), and response inhibition (Stop Signal Task; SST). Determinants of performance on the CANTAB, including age, maternal/child IQ, and sex, were analyzed using Spearman correlation and Welch t-tests. Equivalence testing compared performance with existing norms.

Results: Performance improved with age on all measures. Child-IQ was mildly associated with measures of memory and attention but not executive functioning, consistent with American norms. Maternal-IQ was not associated with any outcomes, and males performed better on IED. Mexican norms were comparable to American norms on almost all outcomes, with the exception of a short-term visual recognition memory task (DMS).

*Corresponding author: Christine Till PhD, Department of Psychology, York University, 4700 Keele Street, M3J 1P3 Toronto (ON), Canada ctill@yorku.ca.

Conflict of Interests: The authors declare no conflict of interests.

Conclusions: This study provides the largest normative data for CANTAB youth performance in a community sample of Mexican youth. Findings demonstrate the expected maturational effects of executive function, specifically in cognitive shifting and inhibition. Levels of executive function performance demonstrated by a Mexican sample were consistent with normative values reported in American youth. These findings, as well as expected associations with child IQ, indicate high applicability of CANTAB for Mexican youth in neurobehavioral studies.

Keywords

neuropsychology; assessment; development; frontal lobe; Mexico City

Introduction

Popularity of neuropsychological assessments as a tool of assessing cognitive domains has increased in the past decades (Hofheimer & Lester, 2008; Keefe, 1995). A neuropsychological assessment is often performed with a battery approach, which involves standardized administration of a variety of tests measuring distinct cognitive abilities (Harvey, 2012). The primary purpose of the assessment is to draw interpretations about brain function by evaluating specific behaviors under controlled conditions (Benton, 1994). Neuropsychological assessments can measure performance in a variety of cognitive domains, including memory, attention, and executive functioning (EF), an umbrella term referring to the cognitive processes that underlie goal-oriented behavior (Loe, Chatav, & Alduncin, 2014). These skills, observed across species, show rapid improvements during the early years of life (Zelazo, 2003), and do not mature until adolescence or even early adulthood (Anderson, 2002; Luciana, Conklin, Hooper, & Yarger, 2005). Thus, it is important to understand the development of these cognitive processes across a broad age range, for example throughout childhood and adolescence (Best & Miller, 2010).

It is proposed that cultural and linguistic factors can impact neuropsychological testing performance (Fernández & Abe, 2017; Perez-Arce, 1999), and this is relevant for non-western and minority populations, especially if the test includes language-based stimuli and instructions (Goudsmit et al., 2016; Ostrosky-Solís & Lozano, 2007). Maintaining the meaning, relevance, and purpose of the assessment is important to produce converging findings and generalizations across samples (Lim et al., 2009). In addition to language and cultural interpretation biases, cross-cultural use of neuropsychological tests can be discriminatory if the construct measures differ across cultures, especially if the tests are not standardized for use across cultures using a representative sample (Van de Vijver & Rothmann, 2004). Culturally fair neuropsychological tests and norms are designed to assess a cognitive outcome, such as intelligence or memory, and are free from the effects of cultural and educational background (Weiss, 1980; Reynolds, 2000; Van de Vijver & Rothmann, 2004). For example, CANTAB (Cambridge Neuropsychological Test Automated Battery) is considered a fair neuropsychological tool because it minimizes potential biases by using abstract visual components and relying on nonverbal responses (www.cambridgecognition.com/cantab). In general, CANTAB subtests have simple instructions, and some tasks may require minimal level of verbal communication to be

completed (Luciana & Nelson, 2002). Despite this, care must be considered when using a neuropsychological instrument in a new population in case of potential cultural influences.

The impact of testing automation should also be considered in the context of cultural-fairness of a test. Many neuropsychological batteries apply computerized technologies in their assessments (Luciana & Nelson, 2002). Advantages of computerized tools include its standardized administration, immediate and accessible feedback, and a detailed recording of accuracy and speed (Fray, Robbins, & Sahakian, 1996). In addition, administration of computerized tests do not require as extensive training as required by clinical neuropsychological tests, and limit the interaction between the examiner and the participant, which may prevent biases and human error. However, a potential weakness is the limited interaction between the examiner and examinee. This may result in a lack of oversight and attention, an important aspect of ensuring validity of the assessment (Cernich, Brennana, Barker, & Bleiberg, 2007). Therefore, computerized batteries are often compared against more traditional psychological assessments in order to assess their validity and suitability in various populations and across age ranges.

CANTAB is a type of a neuropsychological computerized battery. It was developed at the University of Cambridge by Sahakian and Robbins (Fray et al., 1996). CANTAB consists of subtests designed to measure brain functions across three cognitive domains: visual memory, attention and functional components of EF, like working memory, inhibition and attention shifting (Luciana & Nelson, 2002).

CANTAB is administered by using a computer with a touch-sensitive screen and a response key that is used for reaction timing (Fray et al., 1996). Instructions for administering the test is found under the CANTAB test administration guide (www.cambridgecognition.com/cantab). The examiner is allowed to orally provide instructions when necessary (Strauss, Sherman, & Spreen, 2006). CANTAB has been widely used with over 1,500 peer-reviewed papers supporting its cognitive assessment process (www.cambridgecognition.com/cantab). It is frequently used in research studies, and increasingly in clinical practice (Smith, Need, Cirulli, Chiba-Falek, & Attix, 2013). Some validated clinical populations with CANTAB assessments include traumatic brain injury (Maillard-Wermelinger et al., 2009), schizophrenia (Levaux et al., 2007), and Tourette's Syndrome (Rasmussen, Soleimani, Carroll, & Hodlevskyy, 2009).

CANTAB has established a large normative data set from over 2,000 studies with participants aged 4–90 years old. Most of the studies, however, are performed with adult samples, specifically the elderly to assess for age-related cognitive decline (Strauss et al., 2006). There are very few studies that provide age-expected data norms for children (De Luca et al., 2003; Luciana & Nelson, 2002). Current normative values for children include samples of western countries, including American children ($n = 235$) (Luciana & Nelson, 2002), and Australian participants (De Luca et al., 2003). However, Western populations only represent a small portion of the world population, and there can be generalized differences in the attentional and cognitive processing in non-western cultures (Kuwabara & Smith, 2013). Considering that CANTAB tests are non-verbal, it is a practical tool for use in countries of various languages and cultures (Luciana & Nelson, 2002). One study applied

CANTAB to Brazilian youth, and demonstrated its use as a valid assessment tool (Roque, Texeira, Zachi, & Ventura, 2011). However, it is important to note that this study consisted of a relatively small sample ($n = 40$). Considering the dearth of large child sample performance on the CANTAB, and its limited use in non-Western countries, more diverse youth assessment data on CANTAB is warranted.

The current study examined the application of CANTAB in a large Mexican City youth population with a wide age range. Mexican children vary culturally and linguistically as compared to American children. The Organisation for Economic Cooperation and Development (OECD), indicates that Mexico's poverty rate has one of the highest child poverty rate of the OECD members, 25% of Mexican children lived in poor households in 2011, as compared to the OECD average of 14%. Though, these estimates may vary within various classes in Mexico, for example in the working-class, like the current sample.

Further, OECD statistics demonstrate that there are technological variances; 58% of Mexican students have one computer in their homes as compared to 94.5% in the United States (OECD, 2015). Regarding education, Mexican children have a lower chance of completing their schooling (OECD, 2017); up to 50% of 25 to 35-year adults have not completed high school education, compared to almost 10% in the United States. All of these cultural and social variables may affect performance on a computerized assessment battery in a Mexican youth as compared to the Western norms.

Objectives

The primary aim of this study was to analyze CANTAB performance in a large Mexican youth sample from Mexico City and evaluate developmental trends across measures of attention, memory, and executive functioning. Determinants of CANTAB performance in the Mexican cohort, including sex, child IQ, and maternal IQ were also examined. Based on the developmental literature of EF and associated functions (Best & Miller, 2010; Roque et al., 2011), we hypothesized that measures of visual and working memory, executive functioning, and attention will improve in a linear fashion with age. We did not hypothesize a strong effect of sex based on the neuropsychological literature (Luciana & Nelson, 2002; Syväoja et al., 2015). We hypothesized that maternal IQ would be a predictor of better performance (Bacharach & Baumeister, 1998; Forns et al., 2012; Ronfani et al., 2015). The American normative child sample concluded that measures of verbal and non-verbal IQ were moderately associated with better task performance (Luciana & Nelson, 2002), and accordingly, we also hypothesized that contemporaneous child IQ would be associated with performance.

A subsidiary aim tested the equivalence of our Mexican norms to a set of established Western norms. Accordingly, we hypothesized that the Mexican cohort will be comparable to the Western developmental norms, considering that the CANTAB does not rely on language (Roque et al., 2011).

Methods

Participants

Participants consisted of individuals in the cohort from an ongoing environmental study called *Early Life Exposure in Mexico to Environmental Toxicants (ELEMENT)*. The primary purpose of ELEMENT, a birth cohort study begun in 1993 and currently operating with grant support provided by investigators in the School of Public Health at the University of Michigan, the Dalla Lana School of Public Health at the University of Toronto, the Centro de Investigación en Nutrición y Salud, Instituto Nacional de Salud Pública in Mexico City, and the Departamento de Neurobiología del Desarrollo, Instituto Nacional de Perinatología in Mexico City, has been to study the effects of environmental exposures on child health outcomes. The participants in this cohort were not included for having been purposefully exposed to a particular environmental toxin, rather the participants were recruited during their routine prenatal visits to learn how everyday environmental exposures may affect health outcomes.

For the purpose of the current study, a sample of 826 children (405 males and 421 females) aged 5 to 15 years ($M = 9.79$, $SD = 2.6$) born to mothers in the ELEMENT study comprised the study sample. Participants' ages were not uniformly distributed across the ages resulting in unequal sample sizes between age groups (e.g., four 5-year-olds and 169 9-year-olds). Mothers ($n = 826$) were on average 26.8 ($SD = 5.27$) years of age when they gave birth (range = 18–44). A portion of the participants ($n = 217$) had a family socioeconomic status (family possession score) value, which was a questionnaire that asked about the availability of certain items in the home and the SES was calculated based on the sum of the points across all items (Huang et al., 2016). The average score was 6.46 points ($SD = 2.48$, range = 1–14.5).

A portion of the mothers ($n = 515$) had IQ testing ($M = 88.63$, $SD = 13.02$), range = 55–143). They had an average of 10.68 ($SD = 2.89$) years of education. Their children ($n = 826$) were an average of 9.79 ($SD = 2.6$) years of age when they completed the CANTAB (range = 5–15). A portion of the children ($n = 316$) had IQ scores and the scores were standardized, and ranged from $z = -2.67$ to 3.17.

Procedure

CANTAB was administered on a computer with a touch screen and a press pad. Before testing began, the examiner demonstrated to the child how the touch screen system worked and, using scripts provided in the CANTAB manual, to prepare the child before each task. Participants were guided through the assessment with voice over instructions in the Spanish language. The administration of the six subtests took about 65 minutes. Further details for the administration times can be found in Table 1.

Measures

Neuropsychological measures: The neuropsychological constructs examined in this research study are selected from CANTAB battery of cognitive assessments

(www.cambridgecognition.com/cantab). Only the following 6 of the total 23 subtests on CANTAB were administered to all participants (summarized in Table 1).

1. **SOC** (Stockings of Cambridge) is a measure of spatial planning memory requiring individuals to use problem-solving strategies to match two sets of stimuli. Participants are shown two displays, each containing coloured balls. Each participant must move the ball in the lower display to copy the pattern shown in the upper display. During the test, participants are asked to make as few moves as possible to match the two patterns. The variable of interest in this subtest is the *number of perfect solutions*. A higher score indicates better performance.
2. **DMS** (Delayed Matching to Sample) measures visual recognition memory. It is a perceptual matching and delayed visual memory test. Participants are presented with an abstract stimulus on a computer screen. Then, after a short varying delay of (0, 2 or 4 seconds) the same abstract stimulus is shown along with a novel alternative. The task is to select the stimulus they have previously seen. The variable of interest in this subtest is the *percentage of correct solutions*. Higher scores indicate better performance.
3. **IED** (Intra- Extra Dimensional Shift) is a test that features visual discrimination and flexibility of attention. The participant starts by viewing two simple color-filled shapes, and must learn which one is correct by touching it. Computerized feedback teaches the participant which stimulus is correct, and after six correct responses, the stimuli are changed. The test increases in difficulty as participants proceed through the stages. The outcome measure of this subtest is number of *stages reached*. Participants progress through the test by satisfying a set criterion of learning at each stage. The higher the stage reached, the better the performance on the subtest.
4. **RVP** (Rapid Visual Information Processing) is a test measuring sustained attention. Participants are requested to detect patterns of number target sequences (For example: 2–4–6). A white box shown in the centre of the screen containing digits from 2–9 in a random order at a rate of 100 digits per minute. Once the participants see the target sequence they must respond by using the press pad as quickly as possible. The outcome measure of this subtest is *A'*. *A'* is the signal detection measure of sensitivity to the target regardless of response tendency. Higher scores indicate better performance.
5. **SST** (Stop Signal Task) is a test measuring an inhibition response that has already been initiated (Eagle et al., 2008). The participants must quickly respond to an arrow stimulus by selecting one of two options, depending on the direction the arrow is pointing at. If the inhibitor (audio tone) is raised, the participant needs to stop reacting to the stimulus and should withhold any response. The variable of interest in this subtest is *stop signal reaction time (SSRT)*. The less time it takes to respond (i.e., faster response) indicates better performance. In this subtest, lower scores demonstrate better performance.

6. **MTS** (Match to Sample Visual Search Task) is a measure of speed/accuracy trade, which assesses the participant's ability to match visual samples. The participant is shown a complex visual pattern in the middle of the screen. After a short delay, a varying number of similar patterns are shown around the edge of the screen. Only one of these patterns matches the one in the centre of the screen. Efficient performance requires the ability to ignore the distractor patterns and to indicate the correct one. The variable of interest in this subtest is *percent correct*. Higher scores indicate better performance.

In accordance with the American normative study sample conducted by Luciana and Nelson (2002), the first three measures (SOC, DMS, IED) were used in our study as they enabled us to perform direct comparisons between the American and Mexican youth cohorts. In addition, the SST, RVP, and MTS, were administered based on their application for use with children in the literature and so we could develop normative values for the Mexican cohort. The following specific outcomes were selected: *reaction time* for SST (Rouge et al., 2011), *A'* for RVP (Fan, Gau, & Chou, 2014; Lenehan, Summers, Saunders, Summers, & Vickers, 2015; Syväoja et al., 2015), and *percent correct* for MTS (Gau & Huang, 2014). We note that the literature usage of such tools guide our analysis since no normative values are currently established for youth.

Measures of IQ: The *Wechsler Adult Intelligence Scale (WAIS)* is a measure of general intelligence and was used to assess maternal IQ in a postpartum visit. It was administered in the Spanish language to mothers in the cohort. The *WAIS* been shown to be validated for a Mexican cohort making it a standardized and reliable tool (Renteria, Li, & Pliskin, 2010).

The *Wechsler Abbreviated Scale of Intelligence (WASI)* is an abbreviated measure of general intelligence that was administered to the Mexican children at the same time period as the CANTAB to estimate child IQ. It was administered in the Spanish language by a trained psychometrist. The *WASI* includes four subtests (Vocabulary, Similarities, Block Design and Matrix Reasoning). This provides estimates of Verbal IQ, Performance IQ and Full- Scale IQ (Wechsler, 1999).

Statistical Analyses

All statistical analyses were performed in R (R Core Team, 2018) and a nominal Type I error rate of .05 was used for all analyses. In order to address our first objective of analyzing the developmental performance of cognitive functioning demands, the age variable was divided into 8 groups because of the wide (and non-uniform) variability in ages: 5 and 6 year olds, 7 year olds, 8 year olds, 9 year olds, 10 year olds, 11 and 12 year olds (note: only one 12-year-old in the sample), 13 year olds, and 14 to 15 year olds. A Welch one-way independent groups ANOVA was run to test differences across ages by comparing the means across each age bin. A Welch ANOVA was used since it is robust to unbalanced group sizes and unequal variances, as well as mild to moderate nonnormality.

In order to assess for linear effects with age, a linear regression was run predicting each performance outcome using age as a predictor. Since we wished to compare the effects of age across each domain of function, we standardized the subtests by converting each mean

to 0 with a standard deviation of 1. This allowed us to compare the relative beta (standardized) coefficients across each subtest. Lastly, we used these standardized scores to perform independent samples *t*-tests, adjusting with the Holm method, to look at age-related changes from different developmental periods: pre-school age, early elementary school age, middle-school age, and adolescence (note that these groupings were distinct from the age groups used in the analysis above). Association of maternal and child IQ was examined using Spearman correlations since there were outliers (low IQ scores) that would have had an undue influence on the results. Sex differences on all of the subtests were assessed with Welch *t*-tests.

In order to test whether the scores of the American normative sample and the Mexican cohort were equivalent, we first created age bins in our sample that matched those of the American sample (5 and 6-year olds, 7 year olds, 8 year olds, 9 and 10 year olds, 11 and 12 year olds). Therefore, the older ages (13 through 15 year olds) are not present since the American sample did not consist of those ages. Comparison of the Mexican sample to the American norms (Luciana & Nelson, 2002) was assessed via a Welch equivalence test using the TOST function from the *equivalence* package in R (Robinson, 2013). Only the subtests that were the same across both studies were used, which are: DMS (*percent correct*), IED (*stages reached*) and SOC (*number of perfect solutions*). Although the American sample had ages 5 and 6 separately, we combined the groups since there were only four 5-year olds in our sample. Thus, we accordingly combined the 5 and 6 years old in the American norms by weighting the means and SDs based on their sample sizes per age bin. Our upper and lower equivalence bounds were determined as half the SD of the normative sample (American sample).

Results

A one-way heteroscedastic Welch ANOVA, with age as the independent variable, revealed significant mean differences between age groups across all six subtests (see Table 2 and Figures 1–3). Results from the linear regression revealed that performance on each subtest improved across age (Table 3). Comparison of the beta coefficients and R^2 values revealed that DMS, SOC, SST, and RVP are the most influenced by age ($\beta = 0.17$ – 0.20 ; $R^2 = 0.19$ – 0.26), whereas IED and MTS show a weaker age effect ($\beta = 0.08$ – 0.09 ; both $R^2 = 0.05$) (see Table 3). Results from the pairwise *t*-tests comparing pre-school age, early elementary school age, middle-school age, and adolescence revealed that performance on DMS, SOC, SST, and RVP are all continuing to develop (i.e., scores significantly increasing across the developmental stages) into adolescence (i.e., up until age 15), whereas performance on IED and MTS did not increase continuously across all age categories, with neither showing significant differences between ages 10–12 and ages 13–15 (see Table 4).

Sex:

A Welch *t*-test was run to determine whether there were sex differences in performance on the subtests of the CANTAB. Results revealed that there were sex differences for two of the subtests: IED with males ($M = 7.99$, $SD = 0.94$) performing significantly better than females ($M = 7.73$, $SD = 0.95$), $t(819) = -3.91$, $p < .001$, $d = 0.27$) and MTS with females ($M =$

94.40, $SD = 5.38$) performing significantly better than males ($M = 93.59$, $SD = 5.74$, $d = 0.15$), $t(813) = 2.09$, $p < .05$. Both of these effect sizes (Cohen's d) can be considered very small.

One-tailed Spearman correlations were run between maternal IQ, as measured by the WAIS, and performance on the six CANTAB subtests. No significant associations between maternal IQ and CANTAB child-performance were revealed (see Table 5).

The association between child IQ as measured by the WASI and CANTAB performance was also analyzed with one-tail Spearman correlations. Results revealed that Full-scale IQ was significantly and modestly associated with SOC, DMS and MTS; Verbal IQ was significantly and modestly associated with SOC, DMS and MTS; and Performance IQ was significantly and modestly associated with SOC and RVP (see Table 6). Tests of differences in correlations across the Mexican and American samples revealed that the American sample had significantly stronger correlations between DMS and Verbal IQ ($p = .02$) and Performance IQ ($p = .01$), and between IED and Verbal IQ ($p = .04$) than the Mexican sample.

Equivalence testing was performed to determine if the Mexican youth performance on SOC, DMS and IED was statistically similar to that of the American youth. Results revealed that overall, Mexican youth appeared to perform comparable to the American children (Supplemental Table 1). Regarding SOC, the Mexican and American children performed equivalently at age 7, however across the other age groups the results were mixed with the Mexican sample performing better in half the age groups. It is notable that most effect sizes were small and we set a highly conservative equivalence bound. On DMS, results were consistent that the Mexican youth performed slightly worse than the American group on all age ranges. On IED, the groups were statistically equivalent at all age groups.

Discussion

We analyzed CANTAB performance in a large Mexico City youth sample ($n = 826$), the largest dataset, to our knowledge, for CANTAB assessment in children. Performance improved with age across all six subtests of the CANTAB (SOC, DMS, IED, SST, MTS, RVP) in a linear fashion, reflecting the development of attention, memory and EF across childhood. Based on literature, there are currently no normative data on the RVP and MTS subtests (visual attention). Our data show significant improvement in performance as children develop, which demonstrates that attentional skills improve over time, allowing for better on-task focus. Our results also show that children reach adult competency (maximum stages completed) on the IED measure at age 13, and on the MTS (maximum percentage correct = 100%) measure by age 10. This suggests that these outcome measures may no longer be suitable to reflect age-related changes past these ages, and other performance variables that these subtests capture (i.e., reaction time) may be more appropriate for older children. The rest of the skills examined in the other subtests are continuing to develop into adolescence (at least age 15).

Sex was not found to influence performance on any of the subtests with the exception for IED, which is a subtest under the EF domain, and MTS, which is a test of short-term memory. In our study, males performed better on rule acquisition tasks (i.e. IED) than females, and females performed slightly better in detecting accurate matches (i.e. MTS). However, it is important to note that the effect sizes of these differences were negligible to small. In the literature of CANTAB with youth (De Luca et al., 2003; Luciana & Nelson, 2002; Rouque et al., 2011), these sex differences are not detected. However, some studies have found sex differences favoring males or females in selected components of EF, including higher order skill acquisition in typically developing children (Anderson, Anderson, Northam, Jacobs, & Catroppa, 2001; Ardila, Rosselli & Puente, 1994; Kolb & Stewart, 1991).

Results of associations between child IQ and CANTAB performance were somewhat consistent with Luciana and Nelson (2002) (Supplemental Table 2). As expected, several measures (DMS, SOC, MTS) were mildly associated with Full Scale-IQ, mainly reflecting a loading with Verbal-IQ, except for SOC which loaded on both Verbal and Performance IQ. Overall, since CANTAB performance was mildly associated with child-IQ, it suggests that CANTAB is measuring unique cognitive constructs.

Maternal intelligence is an important factor affecting cognitive development in early childhood (Forns et al., 2012; Ronfani et al., 2015). In our study, however, there were no associations between maternal intelligence and child performance on the CANTAB, consistent with the literature suggesting that mother's IQ has a higher bearing on child IQ than on child executive functioning and memory skills.

Equivalence testing revealed that Mexican norms are generally comparable to the American norms, with the exception of Mexican youth performing significantly poorer on DMS, a measure of visuo-spatial short-term memory. Considering that age and other demographic variables were associated with performance in expected ways, it indicates that CANTAB is a suitable battery for Mexican youth. The comparison between our data and the data presented by Luciana and Nelson (2002) in the three CANTAB subtests showed similar age-related results, despite the language and cultural differences between our samples. The fact that CANTAB tests are mostly non-verbal enables its administration in different countries with non-English speaking populations.

However, considering that comparing DMS performance across both samples significantly differed, and that the Mexican correlation coefficients between DMS and IED with Verbal-IQ produced significantly weaker correlations than the American sample, this may suggest some cultural bias with regards to English as a second language or with application to Mexican youth more generally. Overall, since performance on all measures improved significantly across age, these results imply that it is a sensitive measure for assessment of developmental neuropsychological changes in childhood.

Limitations:

Although our study composed the largest normative sample for performance on CANTAB in a child population, it is not without limitations. Since this is a birth cohort population, our

sample was not distributed evenly among the ages and resulted in large variations of sizes within each age. Secondly, CANTAB has not been largely validated in children, such as against other neuropsychological measures and measures of executive functioning, and this is an avenue of future research. On the basis of the weak correlation with IQ, we posit that CANTAB subtests assess cognitive functions that are distinct from IQ.

In addition, educational attainment (i.e., current grade level) was not assessed in this sample. Therefore, it is possible that education may have affected performance in the older ages as educational effects may become more pronounced in the teenage years if fewer children advance to higher levels of education in high school. However, it is important to note that the vast majority of the children (98%) in this sample are ages 14 or younger, which correspond to grades in elementary and middle school. Therefore, it is probable that in our sample, educational achievement, while an important predictor of test performance, is highly uniform with age as it is very likely that the children are still in primary school and have not advanced to high school.

It is also important to mention that our sample was comprised of an urban, working class Mexico City sample, where educational attainment and literacy are not comparable to other areas of Mexico, indicating that these results cannot be generalized to all spheres of Mexican society.

Conclusions

The current study adds significantly to the literature, as it is the largest normative sample for CANTAB performance in youth across a wide age range. Considering the expected developmental trends produced in this sample, our study suggests the application of CANTAB as an appropriate battery for assessing urban, working-class, Spanish-speaking samples on neuropsychological constructs. Considering our large normative dataset, our norms can be used for assessing deficits in executive functions in future research studies, especially those with clinical samples.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

This study was supported by U.S. NIH R01ES021446, NIH R01-ES007821, NIEHS/EPA P01ES022844, NIEHS P42-ES05947, NIEHS Center Grant P30ES017885 and the National Institute of Public Health/Ministry of Health of Mexico. The American British Cowdray Hospital provided facilities used for this research. Dr. David Bellinger collaborated on the design and execution of this study's neurobehavioural testing. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the NIEHS, NIH, or the U.S. EPA. MB and BS had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

References

- Anderson P (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology*, 8(2), 71–82. [PubMed: 12638061]

- Anderson VA, Anderson P, Northam E, Jacobs R, & Catroppa C (2001). Development of executive functions through late childhood and adolescence in an Australian Sample. *Developmental Neuropsychology*, 20(1), 385–406. [PubMed: 11827095]
- Ardila A, Rosselli M, & Puente AE (1994). *Critical issues in neuropsychology*. Neuropsychological evaluation of the Spanish speaker New York, NY: Plenum Press.
- Bacharach VR, & Baumeister AA (1998). Effects of Maternal Intelligence, Marital Status, Income, and Home Environment on Cognitive Development of Low Birthweight Infants. *Journal of Pediatric Psychology*, 23(3), 197–205. [PubMed: 9640899]
- Benton A (1994). Neuropsychological Assessment. *Annual Review of Psychology*, 45(1), 1–23.
- Best JR, & Miller PH (2010). A developmental perspective on Executive Function. *Child Development*, 81(6), 1641–1660. [PubMed: 21077853]
- Cambridge Cognition (<http://www.cambridgecognition.com/cantab>)
- Cernich A, Brennana D, Barker L, & Bleiberg J (2007). Sources of error in computerized neuropsychological assessment. *Archives of Clinical Neuropsychology*, 22, 39–48.
- De Luca CR, Wood SJ, Anderson V, Buchanan J-A, Proffitt TM, Mahony K, & Pantelis C (2003). Normative data from the Cantab. I: Development of executive function over the lifespan. *Journal of Clinical and Experimental Neuropsychology*, 25(2), 242–254. [PubMed: 12754681]
- Eagle DM, Baunez C, Hutcheson DM, Lehmann O, Shah AP, & Robbins TW (2008). Stop-signal reaction-time task performance: Role of prefrontal cortex and subthalamic nucleus. *Cerebral Cortex*, 18(1), 178–188. [PubMed: 17517682]
- Early Life Exposures in Mexico to Environmental Toxicants (ELEMENT). (n.d.). Retrieved August 02, 2017, from <https://sph.umich.edu/cehc/research/element.html>
- Fan L-Y, Gau SS-F, & Chou T-L (2014). Neural correlates of inhibitory control and visual processing in youths with attention deficit hyperactivity disorder: a counting Stroop functional MRI study. *Psychological Medicine*, 44(12), 2661–2671. [PubMed: 24451066]
- Fernández AL, & Abe J (2017). Bias in cross-cultural neuropsychological testing: problems and possible solutions. *Culture and Brain*, 6(1), 1–35.
- Forns J, Julvez J, García-Esteban R, Guxens M, Ferrer M, Grellier J, ... Sunyer J (2012). Maternal intelligence-mental health and child neuropsychological development at age 14 months. *Gaceta Sanitaria*, 26(5), 397–404. [PubMed: 22284911]
- Fray PJ, Robbins TW, & Sahakian BJ (1996). Neuropsychiatric applications of CANTAB. *International Journal of Geriatric Psychiatry*, 11, 329–336.
- Gau SS-F, & Huang W-L (2014). Rapid visual information processing as a cognitive endophenotype of attention deficit hyperactivity disorder. *Psychological Medicine*, 44(02), 435–446. [PubMed: 23561037]
- Goudsmit M, Uysal-Bozkir Ö, Parlevliet JL, Campen JPV, Rooij SED, & Schmand B (2016). The Cross-Cultural Dementia Screening (CCD): A new neuropsychological screening instrument for dementia in elderly immigrants. *Journal of Clinical and Experimental Neuropsychology*, 39(2), 163–172. [PubMed: 27501011]
- Harvey PD (2012). Clinical applications of neuropsychological assessment. *Dialogues Clin. Neurosci*, 4, 91–99.
- Hofheimer JA, & Lester BM (2008). *Encyclopedia of infant and early childhood development* Heidelberg: Elsevier, Academic Press.
- Huang S, Hu H, Sánchez BN, Peterson KE, Ettinger AS, Lamadrid-Figueroa H, et al. (2016). Childhood blood lead levels and symptoms of attention deficit hyperactivity disorder (ADHD): a cross-sectional study of Mexican children. *Environmental Health Perspectives*, 124(6), 868–874. [PubMed: 26645203]
- Keefe RS (1995). The contribution of neuropsychology to psychiatry. *American Journal of Psychiatry*, 152(1), 6–15. [PubMed: 7802121]
- Kolb B, & Stewart J (1991). Sex-related differences in dendritic branching of cells in the prefrontal cortex of rats. *Journal of Neuroendocrinology*, 3(1), 95–99. [PubMed: 19215453]
- Kuwabara M, & Smith LB (2013). Cross-cultural differences in cognitive development: Attention to relations and objects. *Journal of Experimental Child Psychology*, 113(1), 20–35.

- Lenahan ME, Summers MJ, Saunders NL, Summers JJ, & Vickers JC (2015). Does the Cambridge Automated Neuropsychological Test Battery (CANTAB) distinguish between cognitive domains in healthy older adults?. *Assessment*, 23(2), 163–172. [PubMed: 25882162]
- Levaux M-NCAB, Potvin S, Sepehry AA, Sablier J, Mendrek A, & Stip E (2007). Computerized assessment of cognition in schizophrenia: Promises and pitfalls of CANTAB. *European Psychiatry*, 22(2), 104–115. [PubMed: 17227707]
- Lim YY, Prang KH, Cysique L, Pietrzak RH, Snyder PJ, & Maruff P (2009). A method for cross-cultural adaptation of a verbal memory assessment. *Behavior Research Methods*, 41(4), 1190–1200. [PubMed: 19897828]
- Loe IM, Chatav M, & Alduncin N (2014). Complementary assessments of executive function in preterm and full-term preschoolers. *Child Neuropsychology*, 21(3), 331–353. [PubMed: 24754400]
- Luciana M, Conklin HM, Hooper CJ, & Yarger RS (2005). The development of nonverbal working memory and executive control processes in adolescents. *Child Development*, 76(3), 697–712. [PubMed: 15892787]
- Luciana M, & Nelson CA (2002). Assessment of neuropsychological function through use of the Cambridge Neuropsychological Testing Automated Battery: Performance in 4- to 12-year-old children. *Developmental Neuropsychology*, 22(3), 595–624. [PubMed: 12661972]
- Maillard-Wermelinger A, Yeates KO, Taylor HG, Rusin J, Bangert B, Dietrich A, ... Wright M (2009). Mild traumatic brain injury and executive functions in school-aged children. *Developmental Neurorehabilitation*, 12(5), 330–341. [PubMed: 20477562]
- OECD (2015). *Students, Computers and Learning: Making the Connection*, PISA, OECD Publishing, Paris.
- OECD (2017). *Education at a Glance 2017: OECD Indicators*, OECD Publishing, Paris.
- Ostrosky-Solís F, & Lozano A (2007). Neuropsychological assessment in Spanish speaking population. In Boyar LS (Ed.), *New psychological tests and testing research; new psychological tests and testing research* (pp. 157–172). New York, NY: Nova Science Publishers.
- Perez-Arce P (1999). The Influence of Culture on Cognition. *Archives of Clinical Neuropsychology*, 14(7), 581–592. [PubMed: 14590572]
- R Core Team (2017). *R: A language and environment for statistical computing* R Foundation for Statistical Computing, Vienna, Austria URL: <https://www.R-project.org/>.
- Rasmussen C, Soleimani M, Carroll A, Hodlevskyy O (2009). Neuropsychological functioning in children with Tourette syndrome (TS). *J. Can. Acad. Child Adolesc. Psychiatry*, 18(4), 307–315. [PubMed: 19881939]
- Renteria L, Li ST, & Pliskin NH (2010). Reliability and Validity of the Spanish language Wechsler Adult Intelligence Scale (3rd Edition) in a Sample of American, urban, Spanish-speaking Hispanics. *The Clinical Neuropsychologist*, 22(3), 455–470.
- Robinson AP (2013) R package ‘equivalence’ Available at <http://cran.r-project.org/web/packages/equivalence/equivalence.pdf> [Verified 24 July 2015]
- Ronfani L, Brumatti LV, Mariuz M, Tognin V, Bin M, Ferluga V, ... Barbone F (2015). The complex interaction between home environment, socioeconomic status, maternal IQ and early child neurocognitive development: A multivariate analysis of data collected in a newborn cohort study. *Plos One*, 10(5).
- Reynolds CR (2000). Methods for detecting and evaluating cultural bias in neuropsychological tests. In Fletcher-Janzen E, Strickland T, & Reynolds CR (Eds.), *Handbook of cross-cultural neuropsychology* New York: Kluwer Academic/Plenum Publishers.
- Roque DT, Teixeira RAA, Zachi EC, & Ventura DF (2011). The use of the Cambridge Neuropsychological Test Automated Battery (CANTAB) in neuropsychological assessment: Application in Brazilian research with control children and adults with neurological disorders. *Psychology & Neuroscience*, 4(2), 255–265.
- Smith PJ, Need AC, Cirulli ET, Chiba-Falek O, & Attix DK (2013). A comparison of the Cambridge Automated Neuropsychological Test Battery (CANTAB) with “traditional” neuropsychological testing instruments. *Journal of Clinical and Experimental Neuropsychology*, 35(3), 319–328. [PubMed: 23444947]

- Strauss E, Sherman EMS, Spreen O (2006). A compendium of neuropsychological tests: administration, norms, and commentary Oxford: Oxford University Press.
- Syväoja HJ, Tammelin TH, Ahonen T, Räsänen P, Tolvanen A, Kankaanpää A, & Kantomaa MT (2015). Internal consistency and stability of the CANTAB neuropsychological test battery in children. *Psychological Assessment*, 27(2), 698–709. [PubMed: 25528164]
- Van De Vijver A, & Rothmann S (2004). Assessment in multicultural groups: The South African case. *SA Journal of Industrial Psychology*, 30(4).
- Wechsler D (1999). Wechsler Abbreviated Scale of Intelligence. PsycTESTS Dataset
- Weiss SC (1980). Culture Fair Intelligence Test and Draw-a-Person scores from a rural Peruvian sample. *The Journal of Social Psychology*, 111(1), 147–148.
- Zelazo PD (2003). The development of executive function in early childhood Blackwell: Boston, Mass.

Figure 1.
Boxplot of performance by age on Stockings of Cambridge

Figure 2.
Boxplot of performance by age on Delayed Matching to Sample

Figure 3.
Boxplot of performance by age on Stop Signal Task

Table 1.

CANTAB Subtests Used in the Study

CANTAB Subtest	Description	Outcome Measure	Cognitive Domain assessed	Approx. administration time
Stocking of Cambridge (SOC)	Test of participant's ability to engage in spatial problem solving tasks.	Number of perfect solutions	Planning/working memory	10 mins
Delayed Matching to Sample (DMS)	A test of perceptual matching immediate and delayed visual memory.	Percent correct	Short-term visual memory	10 mins
Intra-Extra Dimensional Set Shift (IED)	Test of rule acquisition and reversal, requiring visual discrimination, maintenance and shifting.	Stages reached	Rule acquisition, flexibility and compulsivity	7 mins
Rapid Visual Information Processing (RVP)	A test of sustained attention to detect target number sequences	Alpha prime	Processing speed	7 mins
Stop Signal Task (SST)	Respond quickly to a stimulus, except when a stop signal arrives.	Stop-signal reaction time (SSRT)	Inhibition	20 minutes
Match to Simple Visual Search (MTS)	Tests the ability to match visual samples	Percent correct	Attention	9 mins

Table 2.

Results of Welch ANOVA using Age as Predictor

Age in Years										
Subtest	5–6	7	8	9	10	11–12	13	14–15	F	ω ²
SOC: Number of perfect solutions										
n	21	186	84	169	123	37	63	141	22.04***	0.18
M	5.19	5.92	6.23	6.6	6.89	7.24	7.44	8.35		
Range	1–8	2–9	2–10	2–11	3–12	3–10	3–11	3–12		
SD	1.57	1.58	1.83	1.64	1.73	1.88	2.05	1.93		
DMS: Percent correct										
n	21	185	84	169	123	37	64	141	42.20***	0.26
M	61.55	71.66	74.02	79.11	80.89	82.09	85.47	87.18		
Range	32.5–88.75	40–95	42.5–100	57.5–100	42.5–100	60–97.5	52.5–100	60–100		
SD	14.2	11.08	11.73	9.22	8.66	7.81	10.28	7.89		
IED: Stages reached										
n	21	185	84	169	123	37	64	141	6.41***	0.04
M	7.43	7.64	7.68	7.77	8.03	8.00	8.03	8.18		
Range	7–9	7–9	7–9	7–9	7–9	7–9	7–9	7–9		
SD	0.81	0.87	1.1	0.93	0.92	0.97	0.99	0.92		
SST: Stop signal reaction time										
n	18	186	81	167	123	37	62	140	47.37***	0.28
M	404.15	338.88	309.17	300.76	279.78	260.63	234.61	204.09		
Range	224.05–702.25	149.7–627.8	119.17–526.38	121.72–585.67	130.88–629.2	139.78–560.7	124.85–541.23	91.85–435.45		
SD	128.25	88.7	91.21	96.09	104.24	102.02	79.75	61.26		
MTS: Percent correct										
n	13	186	83	169	123	37	64	141	8.94***	0.06
M	89.58	91.75	93.47	94.59	94.8	94.93	94.14	95.19		
Range	79.17–100	64.58–100	72.92–100	70.83–100	79.17–100	81.25–100	75–100	83.33–100		
SD	6.36	6.26	5.95	5.59	4.94	4.77	4.88	4.03		
RVP: Alpha prime										
n	11	185	83	169	123	36	64	141	43.71***	0.27
M	0.77	0.77	0.77	0.78	0.8	0.82	0.84	0.85		
Range	0.64–0.83	0.51–0.89	0.57–0.91	0.56–0.91	0.55–0.92	0.66–0.91	0.73–0.95	0.7–0.99		
SD	0.05	0.06	0.06	0.07	0.06	0.04	0.05	0.05		

p<.001

Table 3.

Linear Regression with Age as Predictor

	Beta	R²	p
DMS	0.20	0.26	<.001
SOC	0.17	0.19	<.001
IED	0.08	0.05	<.001
SST	−0.19	0.23	<.001
MTS	0.09	0.05	<.001
RVP	0.19	0.24	<.001

Table 4.

Cohen's D values for pairwise tests comparing age groups

	5–7 Vs. 8–9	8–9 Vs. 10–12	10–12 Vs. 13–15
DMS	0.61 ***	0.39 ***	0.64 ***
SOC	0.38 ***	0.29 **	0.57 ***
IED	NS	0.30 **	NS
SST	0.44 ***	0.29 **	0.72 ***
MTS	0.43 ***	NS	NS
RVP	0.22 *	0.41 ***	0.88 ***

Note. P-values adjusted with Holm method

 $p < .001$ **
 $p < .01$ *
 $p < .05$

Table 5.

Spearman correlations between WAIS in mothers and child CANTAB Scores

CANTAB Subtest	<i>Spearman Rho with WAIS</i>
SOC (Number of perfect solutions)	0.11
DMS (Percent correct)	0.06
IED (Stages reached)	0.04
SST (Stop signal reaction time)	−0.02
MTS (Percent correct)	0.08
RVP (Alpha prime)	0.03

Note: WAIS = Wechsler Adult Intelligence Scale; CANTAB = Cambridge Neuropsychological Testing Automated Battery.

Table 6.

Spearman correlations between WASI and CANTAB Scores

CANTAB Subtest	Full scale-IQ	Verbal-IQ	Performance-IQ
SOC (Number of perfect solutions)	0.26 **	0.18 **	0.22 **
DMS (Percent correct)	0.22 **	0.20 **	0.07
IED (Stages reached)	0.00	-0.04	-0.04
SST (Stop signal reaction time)	-0.05	-0.05	-0.02
MTS (Percent correct)	0.18 **	0.15 **	0.02
RVP (Alpha prime)	0.10	0.05	0.13 *

Note: WASI = Wechsler Abbreviated Scale of Intelligence

*
p < 0.05**
p < 0.01.