The Movement ABC: A Cross-Cultural Comparison of Preschool Children From Hong Kong, Taiwan, and the USA

Susanna M.K. Chow
Hong Kong Polytechnic University

Yung-Wen Hsu
National Cheng Kung University

Sheila E. Henderson
University of London

Anna L. Barnett
Oxford Brookes University, UK

Sing Kai Lo
Deakin University, Australia

The purpose of this study was to evaluate the suitability of the Movement Assessment Battery for Children (M-ABC) for use in Greater China. Chinese children numbering 255 between the ages of 4 and 6 from Hong Kong and 544 from Taiwan were tested individually on the standardized test contained within the M-ABC. Data from these 799 children were compared to that presented in the test manual for the 493 children of the same age comprising the United States standardization sample. Both within-culture and cross-cultural differences were statistically significant when all items of the M-ABC were examined simultaneously, but effect sizes were too low to be considered meaningful. However, descriptive analysis of the cut-off scores used for impairment detection on the test suggested that adjustments to some items would be desirable for these particular Chinese populations.

Many instruments are available for the assessment of movement disorders in children (Barnett & Peters, 2004; Burton & Miller, 1998). Among standardized instruments designed to measure mild to moderate difficulties, the Movement Assessment Battery for Children (M-ABC; Henderson & Sugden, 1992) is

Susanna M. K. Chow is with the Department of Rehabilitation Sciences, Hong Kong Polytechnic University, Hung Hom, Hong Kong. Yung-Wen Hsu is with the School of Occupational Therapy, College of Medicine, National Cheng Kung University, Tainan, Taiwan. E-mail: yungwen@mail.ncku.edu.tw. Sheila E. Henderson is with the School of Psychology and Human Development, Institute of Education, University of London, UK. Anna L. Barnett is with the Dept of Psychology, Oxford Brookes University, Oxford, UK. Sing Kai Lo is with the Faculty of Health and Behavioural Sciences, Deakin University, Melbourne, Australia.
now recognized as one of the most widely used in the world (Geuze, Jongmans, Schoemaker, & Smits-Engleman, 2001). Studies using the M-ABC have drawn upon many different diagnostic categories. Participants have included children with Attention Deficit Hyperactivity Disorder (for a review, see Harvey & Reid, 2003), Autistic spectrum disorders (for a review, see Smith, 2004), Specific Language Impairment (Hill, Bishop, & Nimmo-Smith, 1998), Congenital Hypothyroidism (Kooistra, Schellekens, Schoemaker, Vulsma, & van der Meere, 1998), epilepsy (Beckung, Uvebrandt, Hedstrom, & Rydenhag, 1994), neurofibromatosis type 1 (North et al., 1994) childhood encephalitis (Rantala, Uhari, Saukkonen, & Sorri, 1991) and hemiplegia (Mercuri et al., 1999) as well as with Developmental Coordination Disorder (DCD).

Whatever the etiology of a child’s movement difficulty, increased awareness of the problems such children face and knowledge of their consequences have replaced the question of whether it is necessary to do something for these children, with a more urgent enquiry as to what we should do (Cermak & Larkin, 2002; Henderson & Henderson, 2002). Linked with questions regarding intervention, of course, must be questions regarding the most appropriate means of assessment. One aspect of the latter debate involves evaluating the suitability of particular instruments for particular cultures. In this context, questions concerning the familiarity of the test items to a specific cultural group, the effects that administration in a different language might have on test performance, the effects of educational experience on children’s responses to being assessed, the effect of the gender of the tester on performance, and many other factors must all be considered as part of a comparison between the cultural group in question and the sample of children comprising the original test standardization sample (Kline, 2000).

Since the M-ABC appeared in 1992, a number of studies reporting comparisons between children in the original standardization sample from the United States of America (USA) and children from European countries have been published (e.g., Rosblad & Gard, 1998; Smits-Engelsman, Henderson, & Michels, 1998). In contrast, studies from the Far East have been few (e.g., Miyahara et al., 1998). Since many therapists and educators in that part of the world are now interested in using the M-ABC as a diagnostic tool, information on cultural appropriateness would be extremely useful.

In 1998, Henderson, the first author of the M-ABC, began to work with colleagues in Hong Kong, Taiwan, and Mainland China on a project designed to evaluate the suitability of the M-ABC for use with children of Chinese ethnic origin. Since early identification and intervention for children with movement difficulties is considered desirable by educators and therapists in most countries, the project began by focusing on children four to six years of age. In our first study, based in Hong Kong (now known as a Special Administrative Region of China), we focused on the standardized test, which is one of two assessment components within the M-ABC. This study showed that the test could be administered reliably to Chinese children in Cantonese and that test-retest reliability was comparable to that reported in the test manual (Chow & Henderson, 2003). In addition, a number of differences between Chinese children from Hong Kong and USA children who formed the 1992 standardization sample were reported (Chow, Henderson, & Barnett, 2001). The stability of these differences and the implications for producing test norms for Chinese children as a whole form the basis of this paper.
Roughly a quarter of the world’s population is Chinese, with many subgroups living both within the Greater China region and overseas. An examination of the similarities and differences between various subsets within this ethnic group adds a new dimension to cross-cultural studies. To the best of our knowledge, this has never been attempted in the motor domain. In the present study, we attempt a comparison between Chinese children from Hong Kong and from Taiwan, countries which share the same ethnic origin but speak different languages and have with very different political and economic histories.

Specifically, the aims of this study were three: (a) to ensure that the M-ABC test could be used reliably by testers speaking in mandarin, the language used in Taiwan, as opposed to Cantonese, used in Hong Kong. To achieve this objective, inter-tester reliability was checked on a sub-set of the children comprising the main sample; (b) to compare the performance of pre-school children from Hong Kong and from Taiwan on the M-ABC; and (c) to compare the Chinese children with their age-peers from the USA.

Method

Participants

The standardization sample of the original M-ABC comprised a total of 1234 children between the ages of 4 and 12 from the USA (Henderson & Sugden, 1992). In the present study, only the data for the 493 children between the ages of 4 and 6 years of age were employed (males 264, females 229). The sampling technique used to select the children from the USA is outlined in the M-ABC manual (p. 199-200). To obtain samples that were as representative as possible of the populations of Hong Kong and Taiwan, we adopted the same main variables as were used in the USA. These were age, gender, geographical location, and socio-economic status (SES). Ethnic origin, however, was not an issue in this project since 99% of the people living in Hong Kong and 98% of those living in Taiwan are descendents of families originally from Mainland China. Since the tasks within the M-ABC lie outside the compass of children with severe cognitive, sensory, and/or physical disabilities, such children have never been included in any standardization sample.

Chow et al. (2001) described in detail the selection and characteristics of the 255 children from Hong Kong. To obtain our Taiwanese sample, we began in Tainan, in the South of the country. Kindergartens and nursery schools were first selected to represent each of the seven districts of the region. Head teachers from participating schools were then asked to select children randomly from each age group until the required balance of boys and girls was obtained. Any child with a severe sensory, physical, or intellectual disability known to the teachers was excluded from the study. Using these criteria, a total of 330 children from this location were selected, 176 boys and 154 girls. As in the USA and in Hong Kong, a measure of socio-economic status (SES) was obtained using the level of parental education as the best estimate. This information was provided for each child by the school and categorized into three levels: no secondary education, secondary education, and post-secondary education. After we had satisfied ourselves that SES and district within the region (see below) were not factors which needed to be considered in our subsequent recruitment of children, the same sampling technique was used to
select 57 children (31 boys, 26 girls) from the city of Taichung in the middle of the country, 87 from Taipei in the north (45 boys and 42 girls), and 70 more from Tainan, making a total of 544 children.

**Instrument**

The M-ABC contains two assessment instruments, a standardized performance test and a criterion-referenced teachers’ observational checklist. The standardized test, which forms the focus of the present study, accommodates four age bands: Age Band One (AB1) is designed to be used with children age 4 through 6 years, AB2 serves 7 and 8 year olds, AB3 serves 9 and 10 year olds, and AB4 serves those of 11 or more years. For each age band, the M-ABC has eight test items grouped under the headings Manual dexterity, Ball skills, and Balance. Table 1 provides a brief description of the test items and the units of measurement for each task. As Posting Coins and One-leg Balance are tested in both the preferred and non-preferred limbs, a total of 10 scores are obtained. Normative data on children from 4 to 12 years of age is provided in the manual. Reliability and validity data can be found in the M-ABC manual and in Barnett and Henderson (1998) as well as in more recent published studies (e.g., Chow & Henderson, 2003; Croce, Horvat, & McCarthy, 2001).

**Procedure Used in Taiwan**

The M-ABC manual has been translated into the Chinese writing system shared by all persons of Chinese origin, whatever their spoken language (Chow, Meng,

**Table 1  Distribution of Children in the Tainan Sample by Age, Gender, and District**

<table>
<thead>
<tr>
<th>Age</th>
<th>East</th>
<th>West</th>
<th>Central</th>
<th>South</th>
<th>North</th>
<th>Anping</th>
<th>Annan</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>22</td>
<td>6</td>
<td>7</td>
<td>14</td>
<td>10</td>
<td>31</td>
<td>15</td>
<td>105</td>
</tr>
<tr>
<td>5</td>
<td>21</td>
<td>6</td>
<td>6</td>
<td>17</td>
<td>12</td>
<td>28</td>
<td>14</td>
<td>104</td>
</tr>
<tr>
<td>6</td>
<td>29</td>
<td>6</td>
<td>8</td>
<td>18</td>
<td>16</td>
<td>33</td>
<td>11</td>
<td>121</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Girl</td>
<td>31</td>
<td>9</td>
<td>9</td>
<td>24</td>
<td>19</td>
<td>45</td>
<td>17</td>
<td>154</td>
</tr>
<tr>
<td>Boy</td>
<td>41</td>
<td>9</td>
<td>12</td>
<td>25</td>
<td>19</td>
<td>47</td>
<td>23</td>
<td>176</td>
</tr>
<tr>
<td>Total</td>
<td>72</td>
<td>18</td>
<td>21</td>
<td>49</td>
<td>38</td>
<td>92</td>
<td>40</td>
<td>330</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Sample %</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample %</td>
<td>27.5</td>
<td>27.4</td>
</tr>
<tr>
<td>Population</td>
<td>4.5</td>
<td>4.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>6.0</th>
<th>6.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample %</td>
<td>14.0</td>
<td>14.0</td>
</tr>
<tr>
<td>Population</td>
<td>24.0</td>
<td>24.0</td>
</tr>
<tr>
<td></td>
<td>13.8</td>
<td>13.8</td>
</tr>
<tr>
<td>Sample %</td>
<td>24.0</td>
<td>24.0</td>
</tr>
<tr>
<td>Population</td>
<td>10.3</td>
<td>10.3</td>
</tr>
</tbody>
</table>
Hsu, & Wu, 2002). In Taiwan, the verbal instructions were delivered in Mandarin, the language spoken in that country (in Hong Kong, the spoken language used was Cantonese). Three final-year occupational therapy students administered the test. All had been trained in the use of the M-ABC by the second author. This involved reading the test manual in both English and Chinese, viewing a video of the test being administered, assessing a number of typically developing children who did not form part of the sample described here, discussing any problems encountered and being closely supervised during the first month of the four-month testing period testing.

In all cases, children were assessed individually in their own preschools. The room used varied from school to school; in some cases, the gymnasium was employed, in others an empty classroom. Depending on the age of the child, the test duration ranged from 20 to 40 minutes. For the intertester reliability component of the study, one assessor tested each child, while the other two observed and scored performance independently. Each of the three assessors took the lead role in the assessment in turn until a total of 25 children had been tested.

### Scoring and Data Analysis

Children’s performance on the M-ABC is recorded in various ways. For every individual item, a record is kept of raw scores, such as the number of catches made or the number of seconds taken to complete a task. As Table 1 indicates, the number of trials administered and the direction of scores vary by task. For example, a high score can be good (many successful catches) or poor (being slow and taking many seconds to complete a task). Under normal circumstances, these scores are then converted into scaled scores that may take a value from 0 to 5 for each item, higher scores always indicating poorer performance. For items that require both limbs to be tested, such as balancing on one leg, scores for each limb are summed and divided by two. This then yields a total of eight scores, which when summed produce a total out of 40. Since all of the scaled scores, both individual and total, are derived from normative samples, new studies that are explicitly designed to examine the appropriateness of the published norms for a particular group of children must take as their starting point the original raw scores.

In the present study, reliability estimates for the raw scores were obtained using ANOVA-based intraclass correlation ICC (2,1; Shrout & Fleiss, 1979; Suen & Ary, 1989). Unlike reliability measures such as Pearson and Spearman correlations, intraclass correlations are not affected by systematic bias. Moreover, ICC values are more easily interpreted than the above mentioned measures in that they allow the partitioning of variance into that attributable to “true” and “random” variation. Thus, an ICC of, say, .8 implies that only 20% of the total variance is attributable to random error. It has also been shown that kappa, one of the most commonly used indices of interrater reliability, is in fact the categorical equivalent of ICC (Fleiss, 1981).

Intragroup differences within the Taiwan sample as well as intergroup differences between the various Chinese and USA groups were examined separately either by MANOVA or MANCOVA, as appropriate, with the 10 M-ABC items being the dependent variables. In view of the large number of comparisons being made, sharpened Bonferroni corrections (Hochberg & Benjamini, 1990) were applied to
all post-hoc tests. Sharpened Bonferroni correction is superior to Bonferroni correction in that the latter typically results in a substantial drop of statistical power, especially when the number of comparisons is large. The Sharpened Bonferroni is conceptually similar to and mathematically an extension of the Bonferroni method except that the power will be higher for a given number of comparisons.

In addition to $p$ values, effect sizes as measured by Eta Squared ($\eta^2$) values were also used for data interpretation. According to Cohen (1988), only $\eta^2$ of $\geq .14$ obtained from MANOVA analyses are considered sufficiently large to be of any consequence.

**Results**

**Reliability of M-ABC Administered in Mandarin**

Table 1 shows the inter-tester reliability data for each test item. ICCs ranged from .93 - .99 with a mean of .96.

**Characteristics of the Taiwanese Sample**

Table 2 shows the numbers of boys and girls at each age in the Tainan sample. As expected, the age, $F(2, 328) = 20.1, p < .001$, $\eta^2 = .39$ and gender of these children, $F(1, 329) = 5.6, p < .001$, $\eta^2 = .15$ was found to be significantly associated with M-ABC scores, but the interaction between age and gender was not significant. Since, however, primary interest in these factors lay in their possible interaction with cultural group, further investigation of these effects was delayed until after the influence of other demographic variables on test performance had been examined. In the next two analyses, therefore, age and gender were treated as covariates.

As Table 2 shows, the proportion of children from each of the 7 districts in Tainan closely resembled the proportions in the population of the region as a whole. For these 330 children, data on SES was also available. A MANCOVA, with age and gender as covariates, was undertaken to investigate the effects of District and SES on test performance. This revealed no effects of either variable (maximum $F = 2.87$; maximum $\eta^2 = .06$), allowing us to combine the data from the 7 districts of Tainan in the north of the country and compare it with that from the other locations, Taichung and Tapei, further south. Since the addition of the other two data sets introduced another factor, which might have influenced test scores, i.e., whether a child lived in an urban or rural setting, the data from all three regions was also dichotomized on this variable and included in the next analysis. Preparation for this analysis showed that the proportion of children from urban and rural areas was fairly similar to that in the country as a whole. In our study, 65% of the children lived in urban settings and 35% in rural areas. Official statistics in Taiwan suggest a ratio of 75 : 25. A MANCOVA, in which geographical location and urban/rural living were entered as main variables revealed no statistically significant effects (maximum $F = 3.09$, maximum $\eta^2 = .006$). In our view, these results justified the combination of data from all 3 regions into one large sample of 544 children, of whom 286 were boys and 258 girls, a proportion which closely resembles the Taiwan population statistic of 52% male to 48% female.
Table 2: Age Band One Items From the Movement ABC, Along With Interrater Reliability Estimates

<table>
<thead>
<tr>
<th>Domain &amp; Item</th>
<th>Number of Trials</th>
<th>Measures</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manual Dexterity</td>
<td></td>
<td></td>
<td>.97</td>
</tr>
<tr>
<td>Item 1 Posting Coins Preferred Hand</td>
<td>Two for each hand</td>
<td>Time taken in seconds</td>
<td>.98</td>
</tr>
<tr>
<td>Item 2 Posting Coins Non-Preferred Hand</td>
<td>Two for each hand</td>
<td>Time taken in seconds</td>
<td>.99</td>
</tr>
<tr>
<td>Item 3 Threading Beads</td>
<td>Two</td>
<td>Time taken in seconds</td>
<td>.93</td>
</tr>
<tr>
<td>Item 4 Bicycle Trail - tracing between 2 lines, 4 mm apart</td>
<td>Two</td>
<td>Number of times the boundaries were crossed</td>
<td>.93</td>
</tr>
<tr>
<td>Ball Skill</td>
<td></td>
<td></td>
<td>.93</td>
</tr>
<tr>
<td>Item 5 Catching Bean Bag with both hands</td>
<td>Ten</td>
<td>Number of successful catches</td>
<td>.93</td>
</tr>
<tr>
<td>Item 6 Rolling Ball through a goal, 40 cm in width</td>
<td>Ten</td>
<td>Number of successful goals</td>
<td>.96</td>
</tr>
<tr>
<td>Item 7 One-leg Balance Preferred Leg</td>
<td>Two for each leg</td>
<td>Duration in seconds (to a max of 20)</td>
<td>.99</td>
</tr>
<tr>
<td>Item 8 One-leg balance Non-Preferred Leg</td>
<td>Three</td>
<td>Pass on 1st, 2nd or 3rd trial</td>
<td>.93</td>
</tr>
<tr>
<td>Item 9 Jumping Over a Cord at knee height</td>
<td>Three</td>
<td>Number of correct steps (to a max of 15)</td>
<td>.93</td>
</tr>
<tr>
<td>Item 10 Walking along a 4.5 metre line with heels raised</td>
<td>Three</td>
<td>Time taken in seconds</td>
<td>.93</td>
</tr>
</tbody>
</table>
Comparison of the two Groups of Children of Chinese Ethnic Origin

Chow et al. (2001) showed that the distribution of children by age, gender, area, and parental education in Hong Kong closely resembled that of the region as a whole. In the present study, therefore, the 255 children from Hong Kong are treated as one representative sample for comparison with the 544 Taiwanese children. Table 3 presents the means and standard deviations of the children’s scores on each test item, averaged across gender and age, for the two samples. In order to explore the possibility that differences in life style, schooling, etc. between the two countries might affect performance on the M-ABC, a MANCOVA was performed on scores shown in Table 3 with age and gender used as covariates. This analysis revealed a statistically significant effect of country, $F(10, 786) = 7.13, p < .001$, but an overall effect size of only .08. When we then examined the 10 item scores individually, two differences reached statistical significance, threading beads ($p < .001$) and catching a beanbag ($p < .001$) with associated $\eta^2$ of .03 and .01, respectively. Overall, therefore, we felt the difference between the HK and TW samples was small enough to warrant combining all children of Chinese ethnic origin.

Comparison of Chinese and USA Children

Table 4 presents the means and standard deviations of the children’s scores on each test item for 1292 children between the ages of the ages of 4 and 6, 799 Chinese children and 493 from the USA. Table 5 summarizes the ANOVA on these data with accompanying effect sizes.

Results of the MANOVA show that the main effects of age ($F = 41.01, \eta^2 = .25$), gender ($F = 13.04, \eta^2 = .09$), and cultural group ($F = 27.90, \eta^2 = .18$) were all statistically significant. However, $\eta^2$ exceeded the recommended value of 0.14 for age and cultural group only. Further univariate analyses of the main effect of age showed that this held for all 10 measures, with older children performing better, faster, more accurately, etc. than younger children ($\eta^2$ exceeded .14 on 6 of the 10 measures). On the remaining four measures, age trends were evident but the differences between some ages was small. With regard to cultural differences, all 10 comparisons reached statistical significance, but none of the $\eta^2$ exceeded .14. Further analysis of the interaction between age and cultural group yielded two statistically significant effects, on the items labeled “bicycle trial” and “walking heels raised.” The $\eta^2$ associated with these effects were .02 and .01, respectively.

Discussion

Awareness of the effects of movement difficulties on children’s lives has increased considerably over the last decade. As one of a range of diagnostic instruments used in the assessment of such difficulties, the M-ABC is well established in Europe and North America. In contrast, therapists and educators in the Far East are only beginning to investigate its suitability (e.g., Miyahara et al., 1998). In a previous study, we considered the suitability of the M-ABC for use with Chinese children in Hong Kong (Chow et al., 2001; Chow & Henderson, 2003). In the present study,
our interest shifts to Taiwan, a first objective being to determine whether a change in the language used for test administration would affect intertester reliability. In fact, no practical difficulties with test administration were revealed. The average intertester reliability was .96, well above Portney and Watkins' (2000) criterion of > .75, for “good reliability.” Since the manual of the M-ABC (Henderson & Sugden, 1992) largely confines itself to reliability and validity data for children of Caucasian origin, these data strongly support the idea that reliability need not be impaired when the test instructions are translated into another language. One possible reason for this robustness is the lack of prescribed verbal instructions for administering the M-ABC. Instead, task demonstration is compulsory and the tester is free to ensure that the child understands the task requirements in whatever way seems appropriate.

Our second objective was to conduct an intracultural comparison between Chinese children living in Hong Kong and in Taiwan. Before doing this, however, it was important to consider the extent to which the samples in the present study were representative of Taiwan as a whole. In Hong Kong (population 6.5 million), resources had allowed us to select a sample of children, albeit small, which met the requirements of representativeness. In Taiwan, with a much larger population (22 million) spread over a more varied landmass, this was not possible. Fortunately, however, we had at our disposal sufficient data from children living within and around the three main cities of Taiwan to conclude tentatively that our samples

Table 3  Means and SDs of Children From Hong Kong and Taiwan Averaged Across Age and Gender

<table>
<thead>
<tr>
<th>Item</th>
<th>HK (n = 255)</th>
<th>TW (n = 544)</th>
<th>F</th>
<th>p</th>
<th>$\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>17.86 (2.92)</td>
<td>18.05 (3.02)</td>
<td>.76</td>
<td>.38</td>
<td>.001</td>
</tr>
<tr>
<td>2</td>
<td>20.57 (3.89)</td>
<td>20.58 (3.79)</td>
<td>.01</td>
<td>.94</td>
<td>.000</td>
</tr>
<tr>
<td>3</td>
<td>53.17 (17.98)</td>
<td>58.92 (18.64)</td>
<td>23.38</td>
<td>&lt;.001*</td>
<td>.029</td>
</tr>
<tr>
<td>4</td>
<td>.49 (.98)</td>
<td>.66 (1.41)</td>
<td>3.14</td>
<td>.08</td>
<td>.004</td>
</tr>
<tr>
<td>5</td>
<td>6.82 (2.33)</td>
<td>7.39 (2.31)</td>
<td>11.29</td>
<td>.001*</td>
<td>.014</td>
</tr>
<tr>
<td>6</td>
<td>6.36 (1.97)</td>
<td>6.10 (2.00)</td>
<td>3.98</td>
<td>.05</td>
<td>.005</td>
</tr>
<tr>
<td>7</td>
<td>14.49 (6.69)</td>
<td>15.27 (5.96)</td>
<td>4.25</td>
<td>.04</td>
<td>.005</td>
</tr>
<tr>
<td>8</td>
<td>13.77 (6.74)</td>
<td>14.73 (6.16)</td>
<td>6.30</td>
<td>.01</td>
<td>.008</td>
</tr>
<tr>
<td>9</td>
<td>1.36 (.85)</td>
<td>1.23 (.68)</td>
<td>5.88</td>
<td>.02</td>
<td>.007</td>
</tr>
<tr>
<td>10</td>
<td>13.87 (2.55)</td>
<td>14.26 (2.17)</td>
<td>5.35</td>
<td>.02</td>
<td>.007</td>
</tr>
</tbody>
</table>

Note. SD = standard deviation, HK = Hong Kong, TW = Taiwan.
*p significant after Sharpened Bonferroni correction
<table>
<thead>
<tr>
<th>Item</th>
<th>4 year olds</th>
<th></th>
<th></th>
<th>5 year olds</th>
<th></th>
<th></th>
<th>6 year olds</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Girls</td>
<td>Boys</td>
<td>Girls</td>
<td>Boys</td>
<td>Girls</td>
<td>Boys</td>
<td>Girls</td>
<td>Boys</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CH</td>
<td>USA</td>
<td>CH</td>
<td>USA</td>
<td>CH</td>
<td>USA</td>
<td>CH</td>
<td>USA</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>20.42</td>
<td>20.61</td>
<td>20.35</td>
<td>22.69</td>
<td>17.22</td>
<td>18.87</td>
<td>17.88</td>
<td>19.78</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.68)</td>
<td>(5.73)</td>
<td>(2.92)</td>
<td>(8.54)</td>
<td>(2.13)</td>
<td>(2.33)</td>
<td>(2.27)</td>
<td>(3.02)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>23.90</td>
<td>24.28</td>
<td>23.18</td>
<td>24.72</td>
<td>20.01</td>
<td>20.85</td>
<td>20.46</td>
<td>21.22</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.57)</td>
<td>(7.72)</td>
<td>(3.75)</td>
<td>(8.67)</td>
<td>(2.71)</td>
<td>(3.27)</td>
<td>(2.91)</td>
<td>(4.50)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>67.84</td>
<td>66.53</td>
<td>73.56</td>
<td>70.04</td>
<td>53.57</td>
<td>46.35</td>
<td>56.60</td>
<td>50.79</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(18.45)</td>
<td>(29.83)</td>
<td>(22.50)</td>
<td>(25.78)</td>
<td>(14.05)</td>
<td>(16.23)</td>
<td>(11.74)</td>
<td>(16.00)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.10</td>
<td>2.42</td>
<td>1.76</td>
<td>3.24</td>
<td>.18</td>
<td>.87</td>
<td>.43</td>
<td>1.60</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.37)</td>
<td>(2.80)</td>
<td>(2.04)</td>
<td>(2.78)</td>
<td>(.53)</td>
<td>(1.64)</td>
<td>(.90)</td>
<td>(2.19)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>6.58</td>
<td>7.07</td>
<td>6.69</td>
<td>7.02</td>
<td>6.42</td>
<td>7.07</td>
<td>6.97</td>
<td>7.74</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.31)</td>
<td>(2.33)</td>
<td>(2.29)</td>
<td>(2.79)</td>
<td>(2.59)</td>
<td>(2.59)</td>
<td>(2.33)</td>
<td>(2.54)</td>
<td></td>
</tr>
</tbody>
</table>

Table 4  Means and SDs for Girls and Boys From Hong Kong (N = 255), Taiwan (N = 544) and the USA (N = 493)
<table>
<thead>
<tr>
<th></th>
<th>5.06</th>
<th>5.18</th>
<th>5.64</th>
<th>5.98</th>
<th>5.82</th>
<th>6.43</th>
<th>6.76</th>
<th>7.23</th>
<th>6.43</th>
<th>6.89</th>
<th>7.35</th>
<th>7.75</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.82)</td>
<td>(2.30)</td>
<td>(1.88)</td>
<td>(2.42)</td>
<td>(1.97)</td>
<td>(1.96)</td>
<td>(1.79)</td>
<td>(2.22)</td>
<td>(1.81)</td>
<td>(2.03)</td>
<td>(1.80)</td>
<td>(1.78)</td>
</tr>
<tr>
<td>7</td>
<td>12.44</td>
<td>9.84</td>
<td>10.29</td>
<td>8.76</td>
<td>16.31</td>
<td>14.77</td>
<td>15.18</td>
<td>14.47</td>
<td>18.48</td>
<td>17.34</td>
<td>17.74</td>
<td>16.35</td>
</tr>
<tr>
<td></td>
<td>(6.44)</td>
<td>(6.50)</td>
<td>(6.50)</td>
<td>(6.12)</td>
<td>(5.27)</td>
<td>(6.11)</td>
<td>(5.75)</td>
<td>(5.75)</td>
<td>(3.82)</td>
<td>(4.65)</td>
<td>(4.58)</td>
<td>(5.27)</td>
</tr>
<tr>
<td></td>
<td>(6.47)</td>
<td>(6.06)</td>
<td>(5.99)</td>
<td>(6.14)</td>
<td>(5.27)</td>
<td>(6.60)</td>
<td>(6.06)</td>
<td>(6.34)</td>
<td>(3.77)</td>
<td>(5.22)</td>
<td>(5.07)</td>
<td>(5.16)</td>
</tr>
<tr>
<td>9</td>
<td>1.40</td>
<td>1.39</td>
<td>1.21</td>
<td>1.25</td>
<td>1.41</td>
<td>1.15</td>
<td>1.27</td>
<td>1.20</td>
<td>1.15</td>
<td>1.06</td>
<td>1.20</td>
<td>1.06</td>
</tr>
<tr>
<td></td>
<td>(.89)</td>
<td>(.72)</td>
<td>(.66)</td>
<td>(.60)</td>
<td>(.90)</td>
<td>(.52)</td>
<td>(.71)</td>
<td>(.55)</td>
<td>(.49)</td>
<td>(.24)</td>
<td>(.67)</td>
<td>(.29)</td>
</tr>
<tr>
<td></td>
<td>(2.60)</td>
<td>(4.21)</td>
<td>(3.15)</td>
<td>(4.32)</td>
<td>(2.15)</td>
<td>(2.94)</td>
<td>(2.17)</td>
<td>(3.80)</td>
<td>(1.46)</td>
<td>(1.74)</td>
<td>(1.39)</td>
<td>(2.16)</td>
</tr>
</tbody>
</table>

*Note.* SD = standard deviation; CH = Chinese, including children from Hong Kong and Taiwan; USA = the United States of America. For clarification of the measures taken, see Table 2.
### Table 5  ANOVA Results for the Chinese and USA Samples

<table>
<thead>
<tr>
<th>Item</th>
<th>Cultural Group (C)</th>
<th>Age (A)</th>
<th>Gender (G)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$F$</td>
<td>$p$</td>
<td>$\eta^2$</td>
</tr>
<tr>
<td>1</td>
<td>35.49</td>
<td>.000*</td>
<td>.027</td>
</tr>
<tr>
<td>2</td>
<td>14.34</td>
<td>.000*</td>
<td>.011</td>
</tr>
<tr>
<td>3</td>
<td>12.30</td>
<td>.000*</td>
<td>.010</td>
</tr>
<tr>
<td>4</td>
<td>93.24</td>
<td>.000*</td>
<td>.068</td>
</tr>
<tr>
<td>5</td>
<td>17.09</td>
<td>.000*</td>
<td>.013</td>
</tr>
<tr>
<td>6</td>
<td>13.36</td>
<td>.000*</td>
<td>.010</td>
</tr>
<tr>
<td>7</td>
<td>18.52</td>
<td>.000*</td>
<td>.014</td>
</tr>
<tr>
<td>8</td>
<td>21.19</td>
<td>.000*</td>
<td>.016</td>
</tr>
<tr>
<td>9</td>
<td>5.08</td>
<td>.024*</td>
<td>.004</td>
</tr>
<tr>
<td>10</td>
<td>50.62</td>
<td>.000*</td>
<td>.038</td>
</tr>
</tbody>
</table>
### C x A

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>p</th>
<th>$\eta^2$</th>
<th></th>
<th>F</th>
<th>p</th>
<th>$\eta^2$</th>
<th></th>
<th>F</th>
<th>p</th>
<th>$\eta^2$</th>
<th></th>
<th>F</th>
<th>p</th>
<th>$\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.41</td>
<td>.244</td>
<td>.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.82</td>
<td>.441</td>
<td>.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.77</td>
<td>.171</td>
<td>.003</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>9.80</td>
<td>.000*</td>
<td>.015</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>.36</td>
<td>.696</td>
<td>.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>.62</td>
<td>.537</td>
<td>.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>.83</td>
<td>.438</td>
<td>.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>.14</td>
<td>.866</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>2.02</td>
<td>.134</td>
<td>.003</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>7.15</td>
<td>.001*</td>
<td>.011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### C x G

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>p</th>
<th>$\eta^2$</th>
<th></th>
<th>F</th>
<th>p</th>
<th>$\eta^2$</th>
<th></th>
<th>F</th>
<th>p</th>
<th>$\eta^2$</th>
<th></th>
<th>F</th>
<th>p</th>
<th>$\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.77</td>
<td>.184</td>
<td>.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.18</td>
<td>.671</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.36</td>
<td>.552</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.06</td>
<td>.304</td>
<td>.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>.05</td>
<td>.820</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>.06</td>
<td>.800</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1.11</td>
<td>.292</td>
<td>.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>6.90</td>
<td>.009</td>
<td>.005</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>.64</td>
<td>.424</td>
<td>.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>95</td>
<td>.330</td>
<td>.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### A x G

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>p</th>
<th>$\eta^2$</th>
<th></th>
<th>F</th>
<th>p</th>
<th>$\eta^2$</th>
<th></th>
<th>F</th>
<th>p</th>
<th>$\eta^2$</th>
<th></th>
<th>F</th>
<th>p</th>
<th>$\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.65</td>
<td>.193</td>
<td>.003</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.58</td>
<td>.563</td>
<td>.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.09</td>
<td>.911</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4.42</td>
<td>.012</td>
<td>.007</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.52</td>
<td>.219</td>
<td>.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>.32</td>
<td>.723</td>
<td>.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>.71</td>
<td>.490</td>
<td>.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1.58</td>
<td>.206</td>
<td>.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>2.11</td>
<td>.122</td>
<td>.003</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1.95</td>
<td>.143</td>
<td>.003</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### C x A x G

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>p</th>
<th>$\eta^2$</th>
<th></th>
<th>F</th>
<th>p</th>
<th>$\eta^2$</th>
<th></th>
<th>F</th>
<th>p</th>
<th>$\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.43</td>
<td>.239</td>
<td>.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.40</td>
<td>.671</td>
<td>.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.42</td>
<td>.661</td>
<td>.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>.46</td>
<td>.635</td>
<td>.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>.10</td>
<td>.905</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>.33</td>
<td>.720</td>
<td>.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>.29</td>
<td>.752</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>.22</td>
<td>.805</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>.76</td>
<td>.470</td>
<td>.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>.43</td>
<td>.653</td>
<td>.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p significant after Sharpened Bonferroni correction
were representative of their age group in the country as a whole. In particular, we could find no evidence to suggest that geographical location, urban/rural living, or SES interacted with test performance. This echoed the findings of our USA standardization study.

All children in our Hong Kong and Taiwanese samples attended preschool. In neither country were we able to collect data on children who did not attend preschool, but national statistics report that this proportion is less than 5% in each case (Hong Kong Census and Statistics Department, 2002; Yang & Tsay, 2002). It is sometimes asserted that preschools in Taiwan are slightly less academically oriented than in Hong Kong. For example, by the age of three, most children in Hong Kong can write simple Chinese characters (Opper, 1992), whereas in Taiwan systematic tuition in handwriting starts later (Soong & Lee, 1989). However, we found no evidence to suggest that any curricular differences of this nature influenced performance on the M-ABC. On the two test items, which did yield statistically significant differences, bead threading and catching a beanbag, the direction of the effects went in opposite directions. Children from Hong Kong were faster at beads threading and Taiwanese children more proficient at catching. Moreover, the amount of variance accounted for on these items by regional differences was less than 1% in each case.

Chinese communities tend to preserve their cultural values and customs wherever they live (Dion & Dion, 1996). Thus, ethnic origin and culture are strongly bound together. In this study, we have shown that children of the same ethnic origin from two separate countries with very different political, economic, and cultural histories differed very little in the scores they obtained on the M-ABC. However, ethnic Chinese children in Taiwan and Hong Kong are likely to be more culturally similar than when either group is compared to Chinese children living in San Francisco’s Chinatown. The question of what a “fair” test of competence in various ethnic minority groups in the USA might be lies outside the scope of this paper. However, we have no hesitation in proposing that ethnic Chinese in Hong Kong and Taiwan could safely share one set of test norms, at least for this age band of the test.

Chow et al. (2003) reported differences between children from Hong Kong and the USA standardization sample. In the present study, the absence of major differences between the same children from Hong Kong and a new group from Taiwan allowed us to increase the power of our cross-cultural comparison by aggregating the data from these two countries. Interestingly enough, when this sample of over 750 children was compared to the standardization sample, the pattern of significance found in Chow et al. (2003) was largely replicated. However, large samples such as these may allow us to detect effects of very little practical import. The small effect sizes generally lead us to this conclusion but there are exceptions, which we consider below.

The M-ABC manual includes separate norms for each age group for each test item. In this study, the effect of age on test performance was significant on all 10 measures taken ($p < .0001$ in all cases). For six of these 10, the effect size exceeded .14, meeting Cohen’s criteria for acceptability. Perhaps not surprisingly, these results confirm the need for separate age related norm tables for Chinese children.

With regard to gender, both this and our previous study (Chow et al., 2001) showed that between the ages of 4 and 6, girls tended to be better than boys on
all items except those involving throwing and catching. This pattern of gender differences resembles that obtained in other studies involving older children (e.g., Causgrove Dunn & Watkinson, 1996; Miyahara et al., 1998). Such differences have led some authors to suggest that separate norms for boys and girls should be provided (e.g., Causgrove Dunn & Watkinson 1996; Larkin & Cermak, 2002). However, in the present study, the maximum effect size associated with gender was .04, suggesting that the differences were not be of great consequence. A rather different argument holds that at the present time, separate gender norms for the M-ABC might be premature. This argument reminds us that a primary purpose of the M-ABC is the identification of children with impaired motor competence. Although there may be well-documented differences between boys and girls on the mean scores obtained on individual items, we do not yet have any reliable published evidence on the shape of the distributions at the critical, lower end. Relevant to this problem is the absence of reliable estimates of the true incidence of DCD by gender (see Cairney, Hay, Brent, Mandigo, & Flouris, 2005 for a discussion). Although most studies suggest that there are more boys with DCD than there are girls, the finding is not universal. At the present time, therefore, using a test with separate norms for boys and girls may simply serve to mask true differences in incidence.

Of particular interest in the present study was the effect of the increased sample size on cross-cultural differences. In the area of manual dexterity, Chow et al. (2003) had revealed that Chinese children were generally faster and more accurate than their USA counterparts. In the present study, Chinese children performed better on “posting coins” and “bicycle trail,” but the associated effect sizes were small (maximum .07). On the bicycle trail item, however, the interaction between age and cultural group was also significant, with a very large proportion of Chinese six-year-olds making no errors at all. At age four, 40% of Chinese children obtained perfect scores. By age six, this had become 96%. In this case, the practical implications of such a ceiling effect outweigh the smallness of the effect size associated with this interaction. In order to improve discrimination among Chinese children, therefore, some adjustment to the task, such as making the width of the trail narrower, is clearly indicated.

In our previous study (Chow et al., 2001) the performance of USA children was significantly better than that of children from Hong Kong on the two tasks included in the “ball skills” section of the test. This continued to hold true when children from Taiwan were added to the Chinese sample. However, once again the associated effect sizes were negligible (.01 in both cases), calling into question the practical significance of the finding. Earlier, we had also been surprised to discover that Chinese children actually performed better than USA children on the balance tasks. Opportunities for gross motor play in Hong Kong, in particular, are very limited and parents do not value proficiency in this area of performance very highly (Opper, 1992; Yip, 1999). In the present study, the superiority of Chinese children was replicated and a ceiling effect on the task requiring children to walk along a line with heels raised resulted in a significant interaction between task and cultural group (92% obtained perfect scores at age six). In spite of the small effect size, therefore, some adjustment to this item will also be required.

In contrast to assessment instruments, which are designed to provide information across the entire ability range, the focus of the M-ABC is on the identification
and description of children with coordination difficulties. This is reflected in the scoring system, whereby items are scored negatively with high scores representing poor performance. For each test item, raw scores are converted to scaled scores and children are assigned a score of 0 through 5. These scores are then summed to produce a total, with a maximum of 40. When interpreting a child’s scores on the test (either at the level of individual item or the total impairment score), emphasis is placed on two regions, those representing degrees of impairment lying between the 5th and 15th percentiles and those below the 5th percentile. In the first instance, the “at risk” concept is invoked and further monitoring suggested. In the second, it is suggested that children with this degree of impairment (below 5th percentile) have a definite movement difficulty. In this study, we evaluated these cut-off points by using the USA norms to determine the proportion of Chinese children falling into these two regions. Since these analyses simply mirrored those obtained by the parametric analyses reported above, we have not reported them in detail. However, two examples might serve to illustrate the issues to be addressed regarding the most suitable norms for a particular population.

Consider first the task of catching a beanbag. Our analyses showed that USA children performed significantly better on this task than did Chinese children, but the effect size was negligible (.01). When the 15th percentile point was considered, we found that 16.4% of Chinese children met this criterion. This is sufficiently close to confirm that little adjustment, if any, is needed. In marked contrast stands the task of posting coins. On this activity, the performance of Chinese children was significantly better with either hand than that of their USA counterparts but the effect sizes remained negligible; however, in this instance, we found that fewer than 5% of Chinese children fell within the important 15th percentile boundary, suggesting that the cross-cultural difference may become more substantial at the lower end of the score distribution, where impairment is at issue. Using the time taken by five-year-olds to post the coins with the preferred hand as an example, the 15th percentile point would have to shift from the > 22 seconds value cited in the test manual to a value of > 19 seconds. While it must be conceded that such effects may merely attest to greater statistical instability at the distributional extremities, these cut-off points lie at the heart of the matter and therefore require especially close scrutiny. In sum, we suggest that the outcome of all three types of analysis, significance tests, tests of effect size, and comparison of proportions around critical cut-points merit consideration in evaluating the need for further normative developments in any new population.

References


**Acknowledgment**

The authors would like to thank Professor Leslie Henderson whose valuable comments helped to improve the content of this manuscript. We would also like to thank all the schools, teachers, and children who participated in the study.