

## **Concurrent Validity of Motor Tests Used to Identify Children With Motor Impairment**

**Swee Kheng Tan, Helen E. Parker, and Dawne Larkin**  
The University of Western Australia

We investigated the concurrent validity and discrimination accuracy of the Bruininks-Oseretsky Test of Motor Proficiency-Short Form (BOTMP-SF) and the McCarron Assessment of Neuromuscular Development (MAND) for identifying children with and without motor impairment (MI). From a total of 69 Australian children aged from 5 to 11 years, 26 children were classified with MI according to three criteria, including the Movement Assessment Battery for Children (MABC), and were age- and gender-matched with 26 non-MI controls. Performance rankings for the MI/non-MI children on BOTMP-SF and MAND tests were highly correlated ( $r_s = .86$ ); however, only 35% of MI cases were classified alike and 71% of cases were agreed on, overall. Comparing each test with MABC, discrimination statistics revealed MAND was the more accurate discriminator of MI, with higher sensitivity and negative predictive values than the BOTMP-SF. The MAND is a more valid test for the identification of MI in Australian children.

Adapted and regular physical educators are increasingly interested in understanding the nature of impaired motor coordination and in accurately identifying the condition. The prevalence of poor coordination in children ranges from 5% to 20% of the population typically served in school physical education classes (Cratty, 1994; Gubbay, 1975; Johnston, Crawford, Short, Smyth, & Moller, 1987; Larkin & Hoare, 1991). Why is it that the prevalence stated by different sources varies so much? Diagnostic criteria and terminology used to describe impaired motor coordination differ and may contribute to variation in identification. See, for example, the Diagnostic and Statistical Manual-IV of the American Psychiatric Association (APA, 1994), ICD-10 Classification of the World Health Organization (World Health, 1996), Gubbay (1975), Henderson and Sugden (1992), and McCarron (1982). One of the more frequently used terms to describe poor motor coordination is developmental coordination disorder (DCD; Polatajko, Fox, & Missiuna, 1995). Clinical diagnostic criteria for DCD require that there is a marked impair-

---

Helen Parker is now at the School of Health and Physical Education, The University of Notre Dame Australia, PO Box 1225, Fremantle, W.A. 6959, Australia. Dawne Larkin is with the Department of Human Movement and Exercise Science, The University of Western Australia, Nedlands, 6907. Swee Keng Tan completed this research as part of her Honors degree under Drs. Parker and Larkin's supervision and is currently a doctoral student at Michigan State University.

ment in the development of motor coordination that significantly interferes with academic achievement and/or performance of activities of daily living, but excludes a diagnosable medical condition such as cerebral palsy, hemiplegia, or muscular dystrophy (American Psychiatric Association, 1994). In this study, we have deliberately used the term *motor impairment* rather than developmental coordination disorder to describe the population of interest. With our physical education background, we believe that poor motor skills, per se, require accurate identification whether or not the impairment adversely impacts the academic performance of the child.

In order that appropriate decisions can be made about intervention, placement, program planning, and performance objectives for children with motor impairment, it is critical that appropriate assessment methods and accurate screening tests be used (Zittel, 1994). Many assessment instruments are readily available for measuring developmental movement skills (see Burton & Miller, 1998, for a recent review). Some of these tests measure a range of motor proficiency from very low to high, for example, the Bruininks-Oseretsky Test of Motor Proficiency (BOTMP; Bruininks, 1978) and the McCarron Assessment of Neuromuscular Development (MAND; McCarron, 1982). Other assessments are specifically designed to identify motor impairment, for example the Test of Motor Impairment-Henderson Revision (TOMI-H; Stott, Moyes, & Henderson, 1984) and its recent revision, the Movement Assessment Battery for Children (MABC; Henderson & Sugden, 1992).

## Tests of Interest

The BOTMP has a long history of use, wide popularity, and recognition in clinical and research applications. It is frequently cited as the motor test of choice in adapted physical education, occupational therapy, and physical therapy research and is used for screening, diagnosis, prescription of treatment, and educational placement decisions (Bruininks, 1978; Miles, Nierengarten, & Nearing, 1988; Riggen, Ulrich, & Ozmun, 1990; Verderber & Payne, 1987; Wilson, Polatajko, Kaplan, & Faris, 1995). The short form comprises a 14-item subset of the 46-item long form and is valid for testing children 3.5 to 14.5 years of age. Initial standardization was based on a sample of 765 U.S. children. Both the long and short forms of BOTMP have been used as the criterion motor assessment test in North America, mainly based on the initial and follow-up validation studies reported by Bruininks (1978) and Broadhead and Bruininks (1983).

The MAND, on the other hand, has much less exposure in the research literature, although it was published in 1982, a few years after BOTMP. The test is designed as a screening, evaluation, and research tool for clinicians, therapists, educators, and researchers (McCarron, 1982). The test comprises 10 fine and gross motor items that assess one- and two-hand dexterity, grip strength, jumping, and balance skills. The items test similar psychomotor abilities as the BOTMP short form, and its test administration time is also comparable. The MAND was standardized on 2,000 U.S. children and provides norms from 3.5 to 18 years. Evidence of the content, construct, predictive, and concurrent validity of the test is provided by McCarron (1982). MAND is currently used for clinical screening in several developmental motor skills programs in Australia and has been used to classify motor proficiency of individuals for experimental research (Hoare, 1994; O'Beirne, Larkin, & Cable, 1994; Rose, Larkin, & Berger, 1994).

The most recently published test of interest to us is the MABC. This test is the latest revision of the TOMI-H with the items in the performance test virtually unchanged from the earlier TOMI-H version but including a new behavioral checklist (Henderson & Sugden, 1992). The MABC serves to screen, identify, and assist in intervention planning and program evaluation. It is also a research tool in the investigation of children with motor impairment. The performance test comprises eight items assessing manual dexterity, ball skills, and static and dynamic balance domains grouped into four age bands covering 4 to 11+ years. Standardization for the MABC was carried out on 1,234 U.S. children. Administration time is similar to both BOTMP-SF and the MAND. The MABC Test's overall reliability and validity are reported to be "good" with much of the evidence for validity drawn from earlier studies with the TOMI-H (Henderson & Sugden, 1992). Its psychometric properties have been assessed across cultures, indicating some difficulties that could lead to future revisions (Kaplan, Wilson, Dewey, & Crawford, 1998; Miyahara et al., 1998; Rosblad & Gard, 1998; Smits-Engelsman, Henderson, & Michels, 1998). Regardless of these limitations, MABC is designed to identify children with motor impairment and is widely used as a criterion assessment tool in research with this population in many countries (e.g., Cantell, Smyth, & Ahonen, 1994; Lefebvre & Reid, 1998; Sigmundsson, Whiting, & Ingvaldsen, 1999; Wilson & Maruff, 1999).

## Test Validity

Test validity is an important issue in the correct identification of children with motor impairment. Validity refers to "the appropriateness, meaningfulness, and usefulness of the specific inferences made from test scores (*Standards for Educational and Psychological Testing*, 1985, p. 9). Only a few researchers have made direct comparisons between motor screening tests. One example is Riggen and colleagues' (1990) evaluation of the concurrent validity between the Test of Motor Impairment - Henderson Revision (TOMI-H) and the Bruininks-Oseretsky - Short Form (BOTMP-SF) for measuring motor skill deficit in 41 children. BOTMP-SF was considered the criterion standard for true identification of motor impairment (MI), and the 15<sup>th</sup> percentile performance standard was the cut off for impaired/nonimpaired classifications in each test. The overall proportion of agreement between impaired/nonimpaired classifications was .88 (88%). Since it exceeded their 80% criterion for satisfactory agreement, Riggen et al. concluded there was a high degree of decision consistency between the tests. However, a closer scrutiny of their MI classification decision revealed a very different picture. In their study, BOTMP-SF identified only 4 MI cases, whereas TOMI-H identified 9, a low 44% MI case agreement. This points to poor decision consistency between the tests for the condition of interest, MI.

In a similar study, Maeland (1992) examined the concurrent validity between the Test of Motor Proficiency (TMP; Gubbay, 1973) and the TOMI-H for identifying motor clumsiness in 223 children. TMP identified only 52.6% of children who were deemed clumsy by TOMI-H (again at the 15<sup>th</sup> percentile standard), and the overall proportion of agreement was a low .51. Even the agreement rate for identifying the group of children with no motor impairment was only 50%. Maeland (1992) therefore concluded there was an inadequate degree of concurrent validity

between TOMI-H and TMP. The most recently published test comparison by Kaplan et al. (1998) also reported a low correlation of .51 between BOTMP-SF and MABC Test and a low Kappa of .33. However, in their study, the sample was drawn from those with learning and attention problems and not primarily motor problems.

Such low levels of agreement for identification of MI are concerning. TOMI-H identified approximately twice as many children with MI as did either BOTMP-SF or TMP. Unidentified motor dysfunction in children may be very distressing as recent research has shown that, without intervention, a child with MI can experience debilitating long term athletic, emotional, physical, social, and academic consequences (Cantell et al., 1994; Geuze & Börger, 1993; Losse et al., 1991). Although most agree with Smyth, Johnston, Short, and Crawford's (1991) plea for early identification and early remedial intervention for motor difficulties in children, identification can be difficult for two reasons. The first reason is that children with DCD or MI, by clinical description, do not have any obvious pathological condition despite persistently displaying significant but subtle movement difficulties (American Psychiatric Association, 1994). Many of these children compensate for or hide their poor motor coordination by avoiding or totally withdrawing from participation in physical activities (Rose et al., 1994; Schoemaker & Kalverboer, 1994). The second reason is that the chosen motor screening tool may have low accuracy and discrimination sensitivity in identifying children with MI, thereby resulting in false negative decisions about the motor status of a child at risk. And despite Keogh, Sugden, Reynard, and Calkins' (1979) call for greater consistency in the measurement and identification of poor coordination in children 20 years ago, the problem of test validity is still an issue. It is important to conduct more research on the validity of motor proficiency tests for identifying children with MI. An assessment tool that incorrectly identifies MI (false positive decision) or fails to identify MI (false negative decision) misrepresents the child's abilities and compromises appropriate case management and remedial intervention (Zittel, 1994).

The research findings of Riggen et al. (1990) and Maeland (1992), although limited in scope, present a disturbing picture as it seems that commonly used screening instruments, at best, have very low agreement for identification of MI. It is clear that more needs to be done to clarify the respective validity and identification accuracy of commonly used tools. Therefore, the purpose of this study was to evaluate the identification accuracy and hence concurrent validity of BOTMP-SF and MAND when the MABC was used as a criterion test for identifying Australian children with motor impairment. This information will provide movement specialists with empirical evidence to guide the choice of valid assessment instruments for identifying children with impaired motor coordination. It will also stimulate further research and development of valid motor skill assessment tools.

## Method

### Participants

An initial convenience sample of 69 participants (44 boys, 25 girls) between the ages of 4 years, 8 months and 10 years, 8 months ( $M = 81.8$  months,  $SD = 19.4$  months) was recruited from a selected city in Australia. Thirty-two of the participants (the referred group) were children on the enrollment waiting list for a move-

ment enhancement program at a university. They were referred with suspected motor coordination problems by teachers, therapists, medical practitioners, or parents. The remaining 37 children (the recruited group) were obtained from surrounding primary schools or the general community in order to provide a comparison group matched for age (within 3 months of birth date) and sex.

Every child included in the group with MI (i.e., the referred group) met the initial two criteria while children in the comparison group did not. The first criterion was referral to a movement enhancement program because of suspected coordination problems. The second criterion was a MABC Test rating at the 15<sup>th</sup> percentile or below as tested by an independent, experienced clinician. All children included in the matched groups met the third criterion, the DCD exclusion criteria set by DSM-IV (APA, 1994) of having no medically diagnosed motor, physical, or psychological problem. Based on these criteria, 26 MI/non-MI pairs were formed ( $N = 52$ ). A summary of the MABC, BOTMP-SF, and MAND scores for the original ( $N = 69$ ) and MI/non-MI groups ( $N = 52$ ) is shown in Table 1. Five children were dropped from the referral group as their scores surpassed the 15<sup>th</sup> percentile cut score for the MABC (range 16<sup>th</sup> to 32<sup>nd</sup> percentile), and the remainder ( $n = 3$ ) failed the additional criteria. Assessments were carried out with the

**Table 1** MABC, BOTMP-SF, and MAND Mean Scores for Initial Sample ( $N = 69$ ) and the Criteria Groups ( $N = 52$ )

Group	<i>N</i>	Age (months)		MABC Impairment Score		BOTMP-SF Standard Score		MAND Neuromuscular Development Index	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Initial	69								
Referred	32			17.9	8.8	49.0	12.2	75.2	17.7
Girls	12	82.2	(22)						
Boys	20	76.3	(16)						
Recruited	37			3.9	3.2	67.2	6.5	105.1	14.5
Girls	13	83.0	(20)						
Boys	24	85.5	(21)						
Criteria	52								
Motor impaired	26			20.1	8.4	47.0	11.8	71.3	17.0
Girls	11	83.4	(22)						
Boys	15	78.7	(17)						
Nonmotor impaired	26			3.7	3.2	68.2	6.6	108.3	14.8
Girls	11	84.5	(21)						
Boys	15	79.5	(17)						

informed consent of both the participants and their parents and with the approval of the university's Human Rights Committee.

## Administration and Test Materials

To initially classify the motor proficiency of the sample of 69 children, a trained, independent examiner administered the MABC Test as the criterion test on Day 1. To avoid rater bias, a second trained examiner, who had no a priori knowledge of the referral background or motor proficiency of any of the children, administered either BOTMP-SF or MAND to each child on that same day. The other test was administered on a second test day ranging from 2–33 days apart. The order of administration of BOTMP-SF and MAND was pseudo-randomly allocated, with half the participants being assessed first on BOTMP-SF, then MAND, and the remainder in the reverse order. All tests took between 15 and 30 min with the time generally increasing as level of performance decreased. Following is a brief description of each test.

The performance test of MABC (Henderson & Sugden, 1992) is comprised of eight items that vary according to the child's chronological age. Four- to six-year-olds (Band 1) perform posting coins, threading beads, bicycle trail, bean bag catch, ball roll, one leg balance, jump over a cord, and walking with heels raised. Seven- and eight-year-olds (Band 2) perform placing pegs, threading beads, flower trail, one hand bounce and catch, throw bean bag, stork balance, jump in squares, and heel-toe walk. Nine and ten-year-olds (Band 3) perform shifting pegs, nut on bolt threading, flower trail, two hand catch, bean bag throw, one board balance, hop in squares, and ball balance. Raw scores on items are converted to Impairment Points, the sum of which (max = 40) is converted to percentile rank. Cut score for MI classification was at the 15<sup>th</sup> percentile based on age-normed total impairment scores (Henderson & Sugden, 1992).

The BOTMP-SF (Bruininks, 1978) comprises two pretest items for determining arm and leg preference and 14 motor skill items sampled from eight subtests. The 14 items are running speed and agility (Subtest 1, 1 item), balance (Subtest 2, 2 items), bilateral coordination (Subtest 3, 2 items), strength (Subtest 4, 1 item), upper-limb coordination (Subtest 5, 2 items), response speed (Subtest 6, 1 item), visual-motor control (Subtest 7, 3 items), and upper-limb speed and dexterity (Subtest 8, 2 items). The standard scores, percentile ranks, and stanine scores were derived from the total item point scores given in the standardized, age-equivalent norms. The cut score for classifying MI was the 15<sup>th</sup> percentile rank, comparable with MABC Test but lower than Bruininks' (1978) identified level of the 22<sup>nd</sup> percentile for low motor proficiency.

The 10-item MAND (McCarron, 1982) assessment includes five fine motor tasks: beads in box (right and left hand), beads on rod (eyes open and closed), finger tapping (right and left hand), nut and bolt (large and small bolt), and rod slide (right and left hand) and five gross motor tasks: hand strength (right and left hand), finger-nose-finger (eyes open and closed), jumping for distance, heel-toe-walk (forward and backward), and standing on one foot (eyes open and eyes closed on each leg). The Neuromuscular Development Index (NDI) was derived from the sum of the 10 item scale scores based on age-equivalent norms with  $M = 100$  and  $SD = 15$ . Again, the 15<sup>th</sup> percentile (equivalent to an NDI score of 84) was chosen as the cut score for MI. This is almost the same as the NDI cut score of 85, recommended by McCarron (1982) for classifying mild MI.

## Statistical Analyses

The degree of concurrent validity between the MABC Test, BOTMP-SF, and MAND was examined using the Spearman rank order correlation on MABC percentiles, BOTMP-SF standard scores, and MAND NDI scores from the initial sample of 69 participants and the MI/non-MI criteria sample of 52. At a more detailed level, decision agreement between MABC, BOTMP-SF, and MAND was measured by calculating the percentage of cases in each of the criteria-selected groups ( $N = 52$ ) that were commonly identified by the tests. Based on these figures, the discrimination accuracy of BOTMP-SF and MAND to correctly identify the 52 criteria cases was determined via three statistics—the test sensitivity (true-positive rate), the test specificity (false-positive rate), and the test predictive ability (Portney & Watkins, 1993). The predictive ability statistic has two parts. The positive predictive value is the proportion of children identified by a test as having MI being actual cases with MI. It is the probability of diagnosing a child with MI who actually has MI. The opposite statistic, the negative predictive value, is the proportion of children identified by the test as non-MI being actual non-MI cases.

## Results

### Validity as Indicated by the Relationship Between Tests

Spearman rank order correlation analyses of performance ranks of all 69 original children across the three tests revealed a correlation coefficient of .79 between the MABC and the BOTMP-SF, a coefficient of .86 between the MABC and the MAND, and a coefficient of .83 between the BOTMP-SF and the MAND. For the smaller subset of 52 matched cases, the respective coefficient values were a little higher with .84 (MABC-BOTMP-SF), .88 (MABC-MAND), and .86 (BOTMP-MAND). These relatively high associations in performance rankings, which accounted for approximately 64 -70% of the common variance, indicate that each test ranked the children's motor performance in a similar way.

### Concurrent Validity: Case Agreement Between the BOTMP-SF and the MAND

The criteria subgroups of 26 children with MI and 26 children without MI were used to establish decision agreement between BOTMP-SF and MAND. Of the 52 criteria cases, BOTMP-SF identified only 8 children with MI, whereas MAND identified 23 (see Table 2). The decision agreement analysis showed that both tests identified the same 8 MI cases and the same 29 non-MI cases. However, the remaining 15 cases were mismatched. They were identified as MI cases by the MAND and non-MI cases by the BOTMP-SF. The resultant overall decision agreement between the two tests was 71%, but only 8 of 23 MI cases were common, a low proportion of agreement (.35).

### Concurrent Validity: Discrimination Accuracy

The discrimination accuracy of BOTMP-SF and MAND compared to MABC Test was assessed by calculating four statistics—the sensitivity, specificity, and posi-

**Table 2** Number of Cases and Decision Agreement Proportions (in parentheses) Between BOTMP-SF and MAND in Identifying Motor Impaired and Nonmotor Impaired Cases (*N* = 52) from the Criteria Sample

		MAND	
		Motor impaired	Nonmotor impaired
BOTMP-SF	Motor impaired	8 (8 of 23 = .35)	0
	Nonmotor impaired	15 (15 of 23 = .65)	29 (29 of 29 = 1.00)

**Table 3** Criteria Sample Cases Identified with MI by MABC and BOTMP-SF

		MABC		<i>n</i>
		Motor impaired	Nonmotor impaired	
BOTMP-SF	Motor impaired	8 <sub><i>a</i></sub>	0 <sub><i>b</i></sub>	8
	Nonmotor impaired	18 <sub><i>c</i></sub>	26 <sub><i>d</i></sub>	44
<i>n</i>		26	26	52

tive and negative predictive values (Portney & Watkins, 1993). The number of cases of MI identified by BOTMP-SF compared to the MI/ non-MI classification from the MABC + criteria (MI *n* = 26 and non-MI *n* = 26) are listed in Table 3.

Discrimination accuracy of the BOTMP-SF was low. Only 8 out of 26 children with MI were identified, resulting in a very low sensitivity of 31% ( $\frac{a}{a+c} = \frac{8}{8+18}$ ). On the contrary, all the children with non-MI were identified (specificity =  $\frac{d}{b+d} = \frac{26}{0+26} = 100\%$ ). A positive predictive value of 100% ( $\frac{a}{a+b} = \frac{8}{8+0}$ ) indicated that all cases positively identified by BOTMP-SF (*n* = 8) were MI cases. The negative predictive value was lower ( $\frac{d}{c+d} = \frac{26}{18+26} = 59\%$ ), indicating that 41% of cases deemed non-MI by BOTMP-SF were in fact MI cases.

Discrimination accuracy for the MAND, reported in Table 4, attained a high 81% ( $\frac{a}{a+c} = \frac{21}{21+5}$ ) for its sensitivity by correctly identifying 21 out of the possible 26 MI cases. However, two non-MI cases were classified as having MI, resulting in a less than perfect specificity value of 92% ( $\frac{d}{b+d} = \frac{24}{2+24}$ ). These two

**Table 4** Criteria Sample Cases Identified with MI by MABC and MAND Tests

		MABC		<i>n</i>
		Motor impaired	Nonmotor impaired	
MAND	Motor impaired	21 <sub>a</sub>	2 <sub>b</sub>	23
	Nonmotor impaired	5 <sub>c</sub>	24 <sub>d</sub>	29
	<i>n</i>	26	26	52

cases scored at the 18<sup>th</sup> and 22<sup>nd</sup> percentiles on the MABC. Of the 23 cases positively identified with MI by MAND, 91% ( $\frac{a}{a+b} = \frac{21}{21+2}$ ) were MI cases. The negative predictive value of 83% ( $\frac{d}{c+d} = \frac{24}{5+24}$ ) indicated that a high percentage of cases identified as non-MI were correctly classified but that 17% of cases deemed to be non-MI by MAND were in fact MI cases.

## Discussion

In this study, we investigated the concurrent validity of the BOTMP-SF and the MAND compared to MABC Test as assessment tools for identifying children, ages 5 to 11, with and without MI. Three different statistical methods were used to evaluate the relative strength of concurrent validity. The first method calculated the Spearman rank order correlation between the MABC, BOTMP-SF, and MAND. The respective coefficients indicated a high association between the tests' rankings of the children's motor performance and implied that the three tests viewed the construct of motor impairment in similar ways. Although this analysis demonstrated that there was high concurrent validity between the tests, further analyses indicated that there were differences in the accuracy with which each test classified the performances into MI or normally coordinated groups.

The second method focused on the decision agreement between tests. The overall decision agreement between the BOTMP-SF and MAND was a moderate 71%, a value that failed to meet Riggen and colleagues' (1990) suggested standard of 80% agreement for a satisfactory level of concurrent validity. This level of agreement was mainly attributable to the perfect agreement in identifying non-MI cases (100%) rather than the agreement for MI cases (35%). The low level of MI agreement can be compared to Riggen et al.'s (1990) finding that BOTMP-SF and TOMI-H agreed on only 4 of 9 MI cases (44%). Such low agreement for MI cases is cause for concern in that a main purpose of motor proficiency screening tests is to identify the majority, if not all, of the children with motor dysfunction from those with regular coordination. An impressive overall percentage agreement conveys false confidence in the concurrent validity of tests if one does not also look at the specific agreement for identification of children with MI. On the basis of decision agreement alone, BOTMP-SF and MAND have relatively low concurrent validity.

The third method of determining concurrent validity examined how accurately the tests discriminate children with MI from those without. The discrimination accuracy statistics of sensitivity, specificity, and positive and negative predictive values provided the information required. The sensitivity statistic conveys the degree to which BOTMP-SF and MAND agree with the criteria-based case classification of the groups. We found that MAND was nearly three times as sensitive as BOTMP-SF when identifying children with MI (81% compared to 31%) and thus concluded that the BOTMP-SF 15th percentile cut score is very conservative in identifying MI cases. Although it has been assumed for a long time that BOTMP-SF is a valid test of motor proficiency by its wide use in many previous studies (for example, Kaplan et al., 1998; Rikken et al., 1990; Spiegel, Steffens, Rynders, & Bruininks, 1990), our results suggest that a 15<sup>th</sup> percentile cut score is too conservative for identification of MI. A discussion of performance cut scores is held below.

The second aspect of discriminative ability is specificity, the degree to which non-MI cases are identified. BOTMP-SF was highly accurate in identifying all children without MI (100%), whereas MAND identified only 92%, because 2 children from the non-MI group were classified as MI. A perfect screening test would obviously be one that has maximum sensitivity and maximum specificity but, according to Portney and Watkins (1993), there are often trade-offs between these two characteristics. Since MAND had a sensitivity ratio of nearly three times BOTMP-SF, despite having a slightly lower specificity, we concluded that MAND was the more accurate motor screening tool. In making this conclusion, we contend that false positive classification decisions, as in the case of MAND misclassifying 8% ( $n = 2$ ) non-MI as MI, are not as significant as the reverse error of false negative classification, that is 69% by BOTMP-SF compared to 19% by MAND. From the viewpoint of identifying children for skills enhancement programs, the classification of children with MI as non-MI has far greater undesirable consequences arising from incorrect decisions about access to intervention than the personal cost of classifying a child with regular coordination as in need of additional movement coaching.

Apart from the sensitivity and specificity, the predictive ability of each test was also calculated to determine its identification accuracy. A positive predictive value represents the proportion of cases classified as MI by a test as being MI cases (Portney & Watkins, 1993). We found that of those identified as MI by BOTMP-SF, all were indeed MI. However, this represented only 8 of the possible 26 MI cases, a number substantially lower than one would expect to be detected from a valid test for identification of MI. By comparison, MAND had a lower positive predictive value of 91% because two non-MI cases were included in the MI group, but this percentage was a detection rate almost three times higher than BOTMP-SF. Although the MAND positive predictive value was lower than that of BOTMP-SF, we contend that this is a relatively small difference when taking into account the great difference in true detection rates between the tests (8 compared to 21 of the possible 26 MI cases). Exclusion from further diagnostic testing or appropriate intervention programs can be detrimental over both the short and long term to the development of a child with MI (Cantell et al., 1994; Geuze & Börger, 1993; Losse et al., 1991).

The final aspect of a test's discriminative ability is negative predictive value. This refers to the probability of those children identified by the test as non-MI

being non-MI cases (Portney & Watkins, 1993). Of all the cases BOTMP-SF classified as non-MI, only 59% were indeed non-MI cases, and a high proportion of MI cases were misclassified as non-MI. This contrasts with MAND classifications that correctly identified 83% of non-MI cases. Again, MAND appears the more accurate discriminator as it minimized the number of non-MI cases that were misclassified.

The degree to which a motor test can discriminate children with MI from those without depends on a number of factors besides content and construct validity and test reliability. These include the cut score standard recommended for clinical classification by a test, the choice of the criterion motor assessment tool, supplementary criteria by which children are deemed to have coordination dysfunction, test items, recency of standardization of norms, and cultural bias in norms.

A critical issue for appropriate classification of performances is the recommended clinical cut score value. Different cut scores result in different sensitivity of detection. In this study, the cut score adopted, the 15<sup>th</sup> percentile, resulted in very low MI detection rates with the BOTMP. However, classification using the less stringent 40<sup>th</sup> percentile identified 1 extra child, and the 50<sup>th</sup> percentile cut off identified 5 extra children with MI (Tan, 1998). These totals are still far fewer than the 26 expected MI cases. So it would seem the cut score level chosen for BOTMP is not responsible for the low detection rate. In other published research with the BOTMP, Lundy-Ekman, Ivry, Keele, and Woollacott (1991) used the 40th percentile rank and Williams, Woollacott, and Ivry, (1992) used the 50th percentile as the cut scores in studies to categorize children into groups with and without movement dysfunction. Experienced users of BOTMP have arbitrarily decided to increase the cut score used for identification of MI, leading to score interpretation difficulties when making effective diagnostic and intervention decisions (Sicol, 1992).

One might question our use of the MABC Test combined with the other criteria for the initial classification of children. It is possible that these criteria for MI classification were perhaps too liberal. However, from Riggen et al.'s (1990) and Maeland's (1992) studies, TOMI-H appears to identify about twice as many children as did the BOTMP-SF and TMP tests, whereas Kaplan et al.'s (1998) study showed relatively weak correlations between BOTMP-SF and MABC for a population with learning and attention problems. If we accept that MABC Test and TOMI-H are really the same motor performance test, then our findings concur with Riggen et al.'s and indicate that MAND is more similar to MABC than is BOTMP-SF from an identification accuracy perspective. But why might this be so? A critique by Hattie and Edwards (1987) of the BOTMP indicated their disquiet about the nature of some of the items included in the test. Some items failed to meet Bruininks' content validity requirement that items represent fine and gross motor skills relevant to everyday tasks—that is, they lacked ecological validity. However, MABC and MAND include items that also fail to link to everyday task performance. Clearly, no published, empirically determined “gold standard” test of motor proficiency with which to compare other tests is recommended.

The recency of test norming and item selection is a relevant issue for test validity. The fact that BOTMP's standardized norms are based on sampling from census data nearly 30 years ago means that current demographic characteristics, societal values about and attitudes toward physical activity, and types of physical activity involvement result in different performance levels in respective age cohorts.

Miyahara et al. (1998) drew on the long history of revisions before the development of the BOTMP to emphasize “. . . the need to continually update item selection and normative data and to relate these to novel cultural contexts” (p. 681). Since the development of the BOTMP in 1978, no update of the test items or norms has been published. And further, since different cultures prefer their own types of physical activities, it cannot be assumed that a test item that is discriminative in a given country and culture will apply across other cultures (Miyahara et al., 1998; Rosblad & Gard, 1998). However, as this criticism could also be directed at MABC and MAND, it does not stand out as a key factor contributing to the very different detection rates for MI by BOTMP-SF and MAND. Cultural bias in test items is a difficult issue to evaluate but should be considered in selection of test batteries. Out-of-date norms and inclusion of culturally inappropriate items would compromise any test’s sensitivity and discriminability and hence validity.

Finally, one might consider the nature of impaired motor coordination. Research has confirmed that it is a heterogenous syndrome (Causgrove Dunn & Watkinson, 1996; Dewey & Kaplan, 1994; Hoare, 1994). Perhaps BOTMP-SF was good at detecting only a particular subtype, whereas MAND was more inclusive of other types that comprise the syndrome. Although the short form and long form of the BOTMP have reportedly good reliability for ranking participants’ performance (Broadhead & Bruininks, 1983; Verderber & Payne, 1987), the discriminability of the subtests within the long form has been questioned by Wilson et al. (1995). The most discriminating four subtest items were running speed and agility, balance, upper limb speed and dexterity, and visual-motor control. In the short form, however, nearly half (6 of 14) items are drawn from the other subtests with low discriminability. Perhaps this selection of items contributed to a much lower detection rate by BOTMP-SF.

In summary, both tests differ in regard to discrimination accuracy. The MAND is, by far, the more sensitive motor screening test for identification of MI. The BOTMP-SF identified children without MI more readily than children with MI and was, therefore, very insensitive to MI. The high sensitivity and negative predictive values point to MAND being the more accurate and valid test of the two, one that correctly identifies as many targeted children as possible. Therefore, although the BOTMP-SF and the MAND showed a high correlation in performance rankings (.83 and .86 for the initial and matched groups, respectively), concurrent validity was not established because there was very low decision agreement for identification of MI. We recommend caution in relying on the BOTMP-SF either as a screening tool in clinical settings or as a group classification test in research. We also question its status as a valid assessment for identifying Australian children with motor impairment.

## References

- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4<sup>th</sup> ed.). Washington, DC: Author.
- Broadhead, G.D., & Bruininks, R.H. (1983). Factor structure consistency in the Bruininks-Oseretsky Test—Short Form. *Rehabilitation Literature*, **44**, 13-18.
- Bruininks, R.H. (1978). *Bruininks-Oseretsky Test of Motor Proficiency examiner’s manual*. Circle Pines, MN: American Guidance Service.

- Burton, A.W., & Miller, D.E. (1998). *Movement skill assessment*. Champaign, IL: Human Kinetics.
- Cantell, M.H., Smyth, M.M., & Ahonen, T.P. (1994). Clumsiness in adolescence: Educational, motor, and social outcomes of motor delay detected at 5 years. *Adapted Physical Activity Quarterly*, **11**, 115-129.
- Causgrove Dunn, J., & Watkinson, E.J. (1996). Problems with identification of children who are physically awkward using the TOMI. *Adapted Physical Activity Quarterly*, **13**, 347-356.
- Cratty, B.J. (1994). *Clumsy child syndromes: Descriptions, evaluation and remediation*. Chur, Switzerland: Harwood Academic Publishers.
- Dewey, D., & Kaplan, B. J. (1994). Subtyping of developmental motor deficits. *Developmental Neuropsychology*, **10**, 265-284.
- Geuze, R., & Börger, H. (1993). Children who are clumsy: Five years later. *Adapted Physical Activity Quarterly*, **10**, 10-21.
- Gubbay, S.S. (1973). A standardised test battery for the assessment of clumsy children. *Proceedings of the Australian Association of Neurologists*, **10**, 19-25.
- Gubbay, S.S. (1975). *The clumsy child: A study of developmental apraxic and agnosic ataxia*. London: W.B. Saunders.
- Hattie, J., & Edwards, H. (1987). A review of the Bruininks-Oseretsky Test of Motor Proficiency. *British Journal of Educational Psychology*, **57**, 104-113.
- Henderson, S.E., & Sugden, D.A. (1992). *Movement Assessment Battery for Children manual*. London: The Psychological Corporation Ltd.
- Hoare, D. (1994). Subtypes of developmental coordination disorder. *Adapted Physical Activity Quarterly*, **11**, 158-169.
- Johnston, O., Crawford, J., Short, H., Smyth, T.R., & Moller, J. (1987). Poor coordination in 5 year olds: A screening test for use in schools. *Australian Paediatric Journal*, **23**, 157-161.
- Kaplan, B.J., Wilson, B.N., Dewey, D., & Crawford, S.G. (1998). DCD may not be a discrete disorder. *Human Movement Science*, **17**, 471-490.
- Keogh, J.F., Sugden, D.A., Reynard, C.L., & Calkins, J.A. (1979). Identification of clumsy children: Comparisons and comments. *Journal of Human Movement Studies*, **5**, 32-41.
- Larkin, D., & Hoare, D. (1991). *Out of Step: Co-ordinating kids' movement*. Perth, Australia: Active Life Foundation, Department of Human Movement, University of Western Australia.
- Lefebvre, C., & Reid, G. (1998). Prediction in ball catching by children with and without a developmental coordination disorder. *Adapted Physical Activity Quarterly*, **15**, 299-315.
- Losse, A., Henderson, S.E., Elliman, D., Hall, D., Knight, E., & Jongmans, M. (1991). Clumsiness in children - Do they grow out of it? A 10-year follow-up study. *Developmental Medicine and Child Neurology*, **33**, 55-68.
- Lundy-Ekman, L., Ivry, R., Keele, S., & Woollacott, M. (1991). Timing and force control deficits in clumsy children. *Journal of Cognitive Neuroscience*, **3**, 367-376.
- Maeland, A.F. (1992). Identification of children with motor coordination problems. *Adapted Physical Activity Quarterly*, **9**, 330-342.
- McCarron, L.T. (1982). *MAND McCarron Assessment of Neuromuscular Development: Fine and gross motor abilities* (rev. ed.). Dallas, TX: Common Market Press.

- Miles, B.H., Nierengarten, M.E., & Nearing, R.J. (1988). A review of the eleven most often-cited assessment instruments used in adapted physical education. *Clinical Kinesiology*, **42**(2), 33-41.
- Miyahara, M., Tsujii, M., Hanai, T., Jongmans, M., Barnett, A., Henderson, S.E., Hori, M., Nakanishi, K., & Kageyama, H. (1998). The Movement Assessment Battery for Children: A preliminary investigation of its usefulness in Japan. *Human Movement Science*, **17**, 679-697.
- O'Beirne, C., Larkin, D., & Cable, T. (1994). Coordination problems and anaerobic performance in children. *Adapted Physical Activity Quarterly*, **11**, 141-149.
- Polatajko, H., Fox, M., & Missiuna, C. (1995). An international consensus on children with developmental coordination disorder. *Canadian Journal of Occupational Therapy*, **62**(1), 3-6.
- Portney, L.G., & Watkins, M.P. (1993). *Foundations of clinical research: Applications to practice*. Norwalk, CT: Appleton & Lange.
- Riggen, K.J., Ulrich, D.A., & Ozmun, J.C. (1990). Reliability and concurrent validity of the Test of Motor Impairment-Henderson Revision. *Adapted Physical Activity Quarterly*, **7**, 249-258.
- Rosblad, B., & Gard, L. (1998). The assessment of children with developmental coordination disorders in Sweden: A preliminary investigation of the suitability of the Movement ABC. *Human Movement Science*, **17**, 711-719.
- Rose, B., Larkin, D., & Berger, B.G. (1994). Perceptions of social support in children of low, moderate and high levels of coordination. *The ACHPER Healthy Lifestyles Journal*, **44**(1), 18-21.
- Schoemaker, M.M., & Kalverboer, A.F. (1994). Social and affective problems of children who are clumsy: How early do they begin? *Adapted Physical Activity Quarterly*, **11**, 130-140.
- Sicoli, F. (1992). Estimating the accuracy of decisions based on cutting scores. *Journal of Psychoeducational Assessment*, **10**, 26-36.
- Sigmundsson, H., Whiting, H.T.A., & Ingvaldsen, R.P. (1999). Proximal versus distal control in proprioceptively guided movements of motor-impaired children. *Behavioural Brain Research*, **106**, 47-54.
- Smits-Engelsman, B.C.M., Henderson, S.E., & Michels, C.G.J. (1998). The assessment of children with developmental coordination disorders in the Netherlands: The relationship between the Movement Assessment Battery for Children and the Körperkoordinations Test für Kinder. *Human Movement Science*, **17**, 699-709.
- Smyth, T.R., Johnston, O., Short, H., & Crawford, J. (1991). The South Australian Screening Test: An evaluation. *Australian Journal of Early Childhood*, **16**(2), 33-37.
- Spiegel, A.N., Steffens, K.M., Rynders, J.E., & Bruininks, R.H. (1990). The Early Motor Profile: Correlation with the Bruininks-Oseretsky Test of Motor Proficiency. *Perceptual and Motor Skills*, **71**, 645-646.
- Standards for educational and psychological testing*. (1985). Washington, DC: American Psychological Association.
- Stott, D.H., Moyes, F.A., & Henderson, S.E. (1984). *The Test of Motor Impairment—Henderson Revision*. Guelph, Ontario: Brook Educational Publishing Ltd.
- Tan, S.K. (1998). *Identifying children with developmental co-ordination disorder using different tests of motor proficiency*. Unpublished honours thesis, University of Western Australia, Nedlands, Australia.

- Verderber, J.M.S., & Payne, V.G. (1987). A comparison of the long and short forms of the Bruininks-Oseretsky Test of Motor Proficiency. *Adapted Physical Activity Quarterly*, **4**, 51-59.
- Williams, H.G., Woollacott, M.H., & Ivry, R. (1992). Timing and motor control in clumsy children. *Journal of Motor Behavior*, **24**(2), 165-172.
- World Health Organization. (1996). *Multiaxial classification of child and adolescent psychiatric disorders*. Cambridge: Cambridge University Press.
- Wilson, B.N., Polatajko, H.J., Kaplan, B.J. & Faris, P. (1995). Use of the Bruininks-Oseretsky Test of Motor Proficiency in occupational therapy. *American Journal of Occupational Therapy*, **49**, 8-17.
- Wilson, P., & Maruff, P. (1999). Deficits in the endogenous control of covert visuospatial attention in children with developmental coordination disorder. *Human Movement Science*, **18**, 421-442.
- Zittel, L.L. (1994). Gross motor assessment of preschool children with special needs: Instrument selection considerations. *Adapted Physical Activity Quarterly*, **11**, 245-260.

## Authors' Notes

We thank Dr. Beth Hands for her professional assistance in conducting the MABC testing and the children and their parents for their willingness to participate in this study.

Correspondence concerning this article should be addressed to Dr. Dawne Larkin, Department of Human Movement and Exercise Science, The University of Western Australia, Nedlands, W.A. 6907, Australia. (Email: dlarkin@cyllene.uwa.edu.au)

The *MAND - McCarron Assessment of Neuromuscular Development: Fine and Gross Motor Abilities* (rev. ed; McCarron, 1982) may be ordered from McCarron-Dial Systems, PO Box 45628, Dallas TX 75245, U.S.A.

The *Movement Assessment Battery for Children* (Henderson & Sugden, 1992) may be ordered from Psychological Corporation Ltd., 24-28 Oval Rd, London NW1 7DX, U.K.

The *Bruininks-Oseretsky Test of Motor Proficiency* (Bruininks, 1978) may be ordered from American Guidance Service Inc., 1401 Woodland Rd, Circle Pines, MN 55014, U.S.A.