Measurement in Physical Education and Exercise Science, 13: 151–165, 2009

ISSN: 1091-367X print / 1532-7841 online
DOI: 10.1080/10913670903048036

Routledge
Taylor & Francis Group

A Field-Based Testing Protocol for Assessing Gross Motor Skills in Preschool Children: The Children's Activity and Movement in Preschool Study Motor Skills Protocol

Harriet G. Williams, Karin A. Pfeiffer, Marsha Dowda, Chevy Jeter, Shaverra Jones, and Russell R. Pate

Department of Exercise Science, Arnold School of Public Health, University of South Carolina at Columbia, Columbia, South Carolina

The purpose of this study was to develop a valid and reliable tool for use in assessing motor skills in preschool children in field-based settings. The development of the Children's Activity and Movement in Preschool Study Motor Skills Protocol included evidence of its reliability and validity for use in field-based environments as part of large epidemiological studies. Following pilot work, 297 children (3–5 years old) from 22 preschools were tested using the final version of the Children's Activity and Movement in Preschool Study Motor Skills Protocol and the Test of Gross Motor Development (2nd Edition). Reliability of the Children's Activity and Movement in Preschool Study Motor Skills Protocol and interobserver reliability were determined using intraclass correlation procedures (intraclass correlation coefficients; ANOVA). Concurrent validity was assessed using Pearson correlation coefficients to compare the Children's Activity and Movement in Preschool Study Motor Skills Protocol to the original Test of Gross Motor Development (2nd Edition). Results indicated that test reliability, interobserver reliability, and validity coefficients were all high, generally above $R/r = .90$. Significant age differences were found. Outcomes indicate that the Children's Activity and Movement in Preschool Study Motor Skills Protocol is an appropriate tool for assessing motor development of 3-, 4-, and 5-year-old children in field-based settings that are consistent with large-scale trials.

Key words: preschool motor skills, assessing motor skills, gross motor skills

INTRODUCTION

Assessment of motor skill development is widely used for determining the overall rate and level of growth and development in early childhood (Burton & Miller, 1998; Bushnell & Boudreau, 1993; Gallahue & Ozmun, 2006). The preschool years are a developmental period during which most children acquire the basic repertoire of locomotor (e.g., running, jumping, galloping, skipping) and object control (e.g., throwing, kicking, catching, striking) skills (Clark, 1994; Wickstrom, 1983; Williams, 1983). They also develop goal-directed motor behaviors and learn

Correspondence should be sent to Dr. Harriet G. Williams, Department of Exercise Science, University of South Carolina, Blatt Center, 101 J, 1300 Wheat St., Columbia, SC 29208. E-mail: hwilliam@mailbox.sc.edu

to combine different movement patterns into sequences to accomplish different goals (Piaget, 1963; Sporns & Edelman, 1993). Motor skill development plays an integral role in the early years when growth, development, and learning frequently center around play and the physical activity associated with it (Clark, 1994, 2005; Williams & Monsma, 2006). Because even young children are now faced with the increasing prevalence of being overweight and inadequate levels of physical activity (Ogden, Carroll, Curtin, McDowell, Tabak, & Flegal, 2006), scientists are examining potential links between motor development and physical activity behavior (Fisher et al., 2005; Graf et al., 2004), and the significant impact the relationship between the two may have on the development of optimal health-related behaviors and future health outcomes (Williams, 2001).

Assessing the quality of the movement processes (e.g., "process characteristics") involved in performing gross motor skills is integral to assessing the overall level of motor skill development in young children (Burton & Miller, 1998; Burton & Rodgerson, 2001; Williams & Monsma, 2006; Wilson, 2005). Assessment of preschoolers' gross motor skills is particularly challenging and important in large-scale studies that involve large numbers of young children. A specific aim of the Children's Activity and Movement in Preschool Study (CHAMPS) was to examine the association between motor skills status and physical activity in children 3 to 5 years of age. A battery that focused on examining critical, qualitative characteristics of motor skill performance (e.g., how the movement is performed) was essential to the CHAMPS project. The CHAMPS Motor Skills Protocol (CMSP), the focus of this article, was developed primarily as a tool for use in diverse field-based settings that are typical of large, multifaceted epidemiological studies. The purpose of this article is to describe the development of the CMSP and provide evidence for its validity and reliability.

METHODS

Initial Development of the CMSP

A team of four professionals, all with extensive experience in motor skill development in young children, was responsible for developing the CMSP. The initial step in developing the CMSP was a review and careful examination of several available and commonly used test batteries (Wiart & Darrah, 2001). Specific tests included, among others, the Ages and Stages Questionnaire (Bricker & Squires, 1999), Movement Assessment Battery for Children (MABC; Henderson & Sugden, 1992), Peabody Developmental Motor Scales (PDMS; Folio & Fewell, 2000), and Test of Gross Motor Development 2nd Edition (TGMD-2; Ulrich, 2000). After careful review of these and other batteries, an expert panel decided on the TGMD-2 as the prototype for developing the CMSP because the concepts embedded in the TGMD-2 were ones that provided a logical and appropriate approach to evaluating young children's motor skill performance. Specifically, the TGMD-2 is one of the few established instruments that uses "process" characteristics of movement performance as a basis for assessing motor skills in young children. Other approaches that might have been used were not as feasible. For example, although the concept of "developmental sequences" was considered as an approach, it was agreed that different aspects of developmental sequences for many skills of interest were not well defined and agreed upon, and that various components of these developmental sequences were difficult to observe without using some external system to record performances for more detailed analysis. The electronic recording of performances, such as videotaping, was not possible given the context of the preschool settings in which we worked (e.g., space limitations, lack of adequate technical capabilities), as well as time constraints and feasibility associated with the large-scale size and nature of the project.

The TGMD-2 was examined closely and administered to 17 preschool children. Performances of the children were observed and scored in situ and with videotape. Based on this preliminary work and the goals of the study, an expert panel determined that the TGMD-2, in its current form, was not appropriate for use in the study for a variety of reasons. For example, process characteristics of several of the motor skills to be observed were not described in concrete behavioral terms, thus making it difficult to observe or determine whether or not these process characteristics were present in children's performances. In addition, the panel determined that certain important process characteristics for a number of skills that were not a part of the TGMD-2 needed to be included. Also, because determining hand/foot preference was important to the study, a standard procedure for doing so was included. It was also important to develop more explicit directions for administering skills and skill sequences.

The expert panel proceeded to develop a framework for the CMSP through discussion and consultation with other professionals. The panel included an early childhood educator, several motor development specialists with experience in administering the TGMD-2, an expert in children's physical activity, and measurement experts. Keeping the purpose of developing a field-based instrument for large-scale studies as a focal point, critical elements to address in developing the protocol were identified. The elements were: (1) inclusion of critical movement process characteristics to assure appropriate assessment of skills, (2) development of behavioral definitions for those process characteristics, (3) inclusion of an appropriate number of trials to assess each skill and promote reliability, (4) demonstration of skills to the children from different perspectives to support their understanding of the skill to be performed, (5) order of administration of skill sequences, (6) establishment of specific guidelines for determining hand/foot preference, and (7) inclusion of an environmental distraction score to a provide a means for describing the environmental setting in which children's performances were observed. Details of these elements are described below.

Number and Nature of Movement Process Characteristics

The panel first identified process characteristics that experts consider integral to identifying differences in the motor skill performance of children of various ages and skill levels. These movement characteristics were examined in relationship to those included in the TGMD-2, and based upon experience with the TGMD-2, decisions were made about which descriptors to include. For example, in the TGMD-2, one process characteristic for the "throw" is described as "hip and shoulder rotation" and is rated as a single action. First, shoulder rotation is difficult to observe, and there were numerous instances in administering the TGMD-2 when children displayed some form of hip rotation but no obvious shoulder rotation. Because variations in "hip–trunk rotation" were more easily observed, hip–trunk rotation was included as one of the movement characteristics of the throw; additional descriptors were added to address the type of rotation occurring at the hip–trunk (e.g., differentiated, block, no rotation). Similar logic was used in defining other process characteristics for various skills. Process characteristics for locomotor skills included in the CMSP are included in Table 1 and for object control skills in Table 2.

TABLE 1
Locomotor Subscale

Skill	*Movement Characteristics*
Run	1. Arms move in opposition to legs, elbows bent 2. Brief period of suspension (both feet off the ground) 3. Narrow foot placement; lands on heel or toe; not flat footed 4. Length of stride even; path of movement horizontal 5. Nonsupport leg flexed to approximately 90° 6. Eyes focused forward
Broad jump	1. Preparatory: flexion of both knees; arms behind body 2. Arms extend forcefully; forward and upward to full extension above the head 3. Take-off and landing on both feet simultaneously 4. Take-off on both feet simultaneously; landing non-simultaneous 5. Arms move downward during landing 6. Balance maintained on landing
Slide	1. Body turned sideways; shoulders aligned with line on floor to initiate 2. Steps sideways with lead foot; slides trail foot next to lead foot 3. Minimum of four continuous step-slide cycles to right 4. Minimum of four continuous step-slide cycles to left 5. Arms used to assist leg action 6. Body maintained in sideways position moving to right 7. Body maintained in sideways position moving to left
Gallop	1. Arms (elbows) flexed and at waist level at take-off 2. Step forward with lead foot; step with trail foot to a position adjacent to or behind lead foot 3. Heel–toe action of lead foot 4. Assumes initial position facing forward 5. Final position facing forward 6. Brief period of suspension; both feet off the floor 7. Maintains rhythmic pattern (four consecutive gallops)
Leap	1. Take off on one foot; land on opposite foot 2. Brief period of suspension (both feet off the ground) 3. Forward reach with arm opposite the lead foot
Hop	1. Nonsupport leg swings forward in pendular motion to assist force production 2. Foot of nonsupport leg remains behind body 3. Arms flexed; swing forward together to produce force 4. Weight received (lands) on ball of foot 5. Takes off and lands three consecutive times on preferred foot 6. Takes off and lands three consecutive times on nonpreferred foot

Number of Trials

The number of process characteristics to be observed for each motor skill was of concern, since the increased number might increase the demands placed on the assessors' observation skills.

TABLE 2
Object Control Subscale

Skill	*Movement Characteristics*
Overarm throw	1. Wind-up initiated by downward movement of hand/arm 2. Hip and shoulder rotated so that nonthrowing side faces target 3. Steps (weight transferred) onto foot opposite throwing arm 4. Differentiated trunk rotation (2) 5. Block trunk rotation (1) 6. Timing of release/flight of ball appropriate (late release = downward flight; early release = upward flight) 7. Arm follows through beyond release (down and across the body)
Underhand roll	1. Ball arm/hand swings down/back of trunk; chest/head face forward 2. Arm action in vertical plane 3. Foot opposite ball hand strides forward toward cones 4. Bends knees; lowers body 5. Ball held in fingertips 6. Ball released close to floor; bounces less than 4 inches high
Kick	1. Rapid and continuous approach to ball 2. Elongated stride or leap immediately prior to ball contact 3. Nonkicking foot placed even with or slightly in back of ball 4. Leg swing is full; full backswing and forward swing of leg 5. Backswing coordinated with forward action of nonkicking leg 6. Ball contacted with instep of kicking foot (shoe-laces) or toe 7. Kicks through ball; leg action does not stop at ball contact
Catch	1. Preparatory: hands in front of body; elbows flexed 2. Arms extend toward ball as it moves closer 3. Ball caught cleanly with hands/fingers (2) 4. Ball trapped against body/chest (1) 5. Ball tracked consistently and close to point of contact 6. Doesn't turn head/close eyes as ball approaches
Stationary strike	1. Dominant hand grips bat just above nondominant hand 2. Nonpreferred side of body faces imaginary "pitcher"; feet parallel 3. Steps (transfers weight) onto foot opposite dominant hand to initiate strike 4. Differentiated trunk rotation (2) 5. Block trunk rotation (1) 6. Arm action/plane of bat movement horizontal 7. Ball contacts bat 8. Swings through ball (action does not stop at ball contact)
Stationary dribble	1. Arm action independent of trunk 2. Ball contacted with one hand at about belt/waist height 3. Pushes ball with fingertips (does not slap at ball with flat hand) 4. Ball contacts surface in front of or to the outside of foot on preferred side 5. Controls ball for four consecutive bounces; feet not moved to retrieve ball

Based on consultation with measurement specialists, four trials of each skill performance were added to avoid difficulties associated with observing a large number of process characteristics to assure that adequate opportunities were provided for observing skill performances and to ensure an appropriate basis for establishing reliability.

Demonstration of Skills

In order to ensure that children better understood what they were being asked to do, two demonstrations were given for each skill. One demonstration was administered with testers facing children; the other was performed with assessors facing the direction in which children were asked to perform the skill. The exception to this procedure was the hop, where both demonstrations were administered with testers facing and moving in the direction in which children were to perform the hop. All demonstrations were given prior to children's performing four trials. No additional demonstrations or feedback were provided by the assessors.

Order of Skill Administration

An order for administration of locomotor and object control skills was established. Based on preliminary work with the TGMD-2, it was noted that children exhibited some confusion in moving from one skill to the next in the order suggested by the TGMD-2. To avoid or minimize this confusion, locomotor and object control skills were administered in different sequences; the following sequence for locomotor skills worked well: run, jump, slide, gallop, leap, and hop. This sequence is a commonly observed, although not universally recognized, sequence in which locomotor skills emerge or are acquired during early development (Williams & Monsma, 2006; Wickstrom, 1983). For object control skills, the order of administration was throw, roll, kick, catch, stationary strike, and dribble. The general guiding principle was the work of Gentile (2000), which suggested that skills in which the object is stationary and in the possession of the individual (e.g., as in the throw or roll) tend to be less complex and thus appear earlier than skills in which the object is moving, involves an implement, and is "outside" the control of the person (e.g., as in dribbling, striking objects).

Hand/Foot Preference

The hop, kick, throw, and roll are skills that involve hand or foot preference. Because hand and foot preferences were of interest, and to assure some consistency in how preference was established, the following guidelines were used to determine preference: (1) hand preference was the hand used in at least three of four trials on both the throw and roll (Bishop, Ross, Daniels, & Bright, 1996; DeAgostini & Dellatolas, 2001; Leconte & Fagard, 2004) and (2) use of the same hand on two or fewer trials on both tasks was an indication of lack of preference or dominance. As with hand preference, the preferred foot was determined by the use of the lead foot in the gallop and the leg used to kick in three of four trials. Use of the same foot on two or fewer trials on both tasks indicated a lack of established foot dominance or preference. A specific procedure for identifying hand and foot preference would allow for the use of dominance and lack of dominance as a potential factor in motor skill development.

Environmental Distraction Score

Children from three types of school settings (i.e., Head Start programs, religious preschools, and commercial childcare centers) were involved in the CHAMPS study. Therefore, motor skills testing was performed within and during a wide range of environmental circumstances and contexts (e.g., hallways, cramped spaces, noisy gyms). Because situational inconsistencies can be an important factor in test administration and interpretation (Doty, McEwen, Parker, & Laskin, 1999; Lyman, 1997), a distraction score was developed. The score was a subjective rating of the nature of the field-based circumstances in which skill observations took place. The measure allowed for the assessment of potential effects of various environmental influences on children's skill performances and the assessors' ability to accurately observe children's skill performances during routine field-based situations.

Four categories formed the basis for the distraction score: (1) noise level (e.g., noise-related disturbances from classrooms, hallways); (2) general distractions (e.g., traffic in the testing area during demonstrations, teachers watching and performing the skill); (3) temperature of the testing area; and (4) nature and adequacy of the space (e.g., inadequate floor space, inappropriate floor surface). Each category was rated on a 1 to 5 Likert-type scale with 1 = *definitely interfering* and 5 = *not interfering*. Testers' ratings from each category were summed to create an overall distraction score that ranged from 4 to 20.

Training and Selection of Motor Development Testers

Because the number of trials and the number and nature of movement process characteristics to be observed might make the assessment process more complex, and because a protocol was being developed, it was determined that testers should have some background in motor development and, if possible, experience working with preschool children. Before administering the CMSP in the actual test setting, extensive training of two testers, who were professionals with a background in motor development, was completed. The testers completed approximately 51 hr of training using both videotape and in situ observations of developmentally diverse preschool children from a pediatric perceptual-motor development program and two private preschools. Children were observed at least twice to develop some consistency in rating process characteristics and to establish an estimate of inter- and intra-observer reliability. Informed consent was obtained from parents and a monetary incentive was provided. Six children from the university-based pediatric motor development program were observed and videotaped. Performances were scored initially while they were videotaped and were also reviewed and scored again later. Five children from a private preschool then were observed, scored, and videotaped as described above. Finally, live observations of six children were conducted at a second private preschool. Two testers were involved—one observed and scored the CMSP while the other administered the motor skill tasks and scored the assessment. These two tester roles were reversed throughout the study and were counterbalanced across children. A professional with extensive experience in young children's motor development also viewed and scored both in situ and videotape observations and served as the "gold standard" criterion during training. In terms of inter-rater reliability, the "gold standard" and testers' scores were significantly correlated (R = .86–.99, $p < .001$) for body control (locomotor), object control, and total scores. Further, intraclass correlation coefficients (ICC; ANOVA) revealed no significant difference among any combination of raters. Intraobserver reliability was also examined during training for data on five participants. Results of ICC analyses indicated that intraobserver reliabilities were high and positive (R = .56–.92).

Preschool Participants

Participants for the main trial included 297 children (158 boys, 139 girls) from 22 preschools. Data were collected on 98 three-year-olds (50 boys, 48 girls), 148 four-year-olds (75 boys, 73 girls), and 51 five-year-olds (33 boys, 18 girls) over an 18-month period (August 2004 through January 2006). Preschools involved were of three types: (1) commercial (n = 11 schools; 136 children), (2) religious-based preschools (n = 7 schools; 88 children), and (3) Head Start programs (n = 4 programs; 73 children). One or more parents of 53% of the children had completed college or technical school or had higher degrees. The remaining 47% of parents either were attending high school or had a high school diploma. The racial make-up of the group was as follows: 51% African American, 40% Caucasian, and 9% Other (Hispanic, Asian American, etc.).

Procedures

The two previously described testers with a background in motor development collected all CMSP field data. Data collection generally lasted for approximately one week per preschool, depending on the number of children from each preschool who were participating and time constraints imposed by the preschool personnel. For most of the testing, both testers were present and administered the assessment protocols. One tester demonstrated motor skills and administered the CMSP, while the other one recorded performance data and rated environmental distractions.

Children were tested individually; the time required was approximately 45 min per child. Testers chose the best possible circumstances for testing in each preschool. Differences within preschool types and among the 22 preschools resulted in a wide variation in testing circumstances and environments. Whenever possible, a gym or long hallway that was free of traffic and distraction was used for assessing children. In some cases, it was not possible to avoid a central location in a preschool and distractions such as teachers or children interrupting assessments occurred occasionally during testing. Children completed the CMSP and the TGMD-2 during the one-week testing period. Appropriate breaks were given as needed to avoid fatigue and maintain attention. Procedures for administering the TGMD-2 were those recommended in the TGMD-2 Manual (Ulrich, 2000). There were two trials per each skill; all process characteristics were rated as 0 or 1. The testers took detailed notes of each testing; the CMSP and TGMD-2 outcome data were entered into a software package (Epi Info™, Center for Disease Control and Prevention (CDC), Atlanta, GA) within one week of testing.

CMSP Outcome Measures

Scores on the CMSP were ratings of movement process characteristics of skills in the Locomotor subscale, Object Control subscale, and Total Test categories. Movement process characteristics were rated as 1 (*present*) or 0 (*not present*) for most skills and summed (separately for 2 and 4 trials) to arrive at a score for locomotor, object control, and total test performances. Exceptions to the ratings of 1 and 0 were made for one process characteristic each of the throw, strike, and hop. For the throw and strike, hip–trunk rotation was scored as 2 (*differentiated*), 1 (*block*), and 0 (*no rotation*). For the hop, a moving hop was scored 2, a stationary hop as 1, and no hop as 0.

These data were used in a series of analyses to examine age differences in gross motor development and to examine validity, test reliability, and interobserver reliability.

Validity

Concurrent validity. To assess the suitability of using the CMSP, performances on the CMSP test and on a criterion test, the TGMD-2 (Wood, 1989), were examined. Performances on both tests were observed during the same time frame. Because the validity of the TGMD-2 has been well established and it is used frequently, the Pearson correlation coefficients were used to compare the CMSP to the TGMD-2.

Construct validity. Construct validity is the degree to which a test measures the construct it was designed to measure (Wood, 1989). Several factors may be used to demonstrate construct validity. Two constructs were examined: age differentiation and correlations among subtests. Because gross motor skill performance is developmental in nature, scores on the CMSP should be strongly associated with chronological age and should differentiate among age groups. With regard to locomotor and object control skills subscale correlations, because both subtests purport to measure gross motor development but from different dimensions, each subscale should have a strong positive correlation with the total test score and a more moderate correlation with each other than with the total test score. A series of analyses were performed using Proc Mixed (controlling for school as a random effect) in SAS (Version 8.2; SAS Institute, Cary, NC) to examine age-related differences. Pearson correlations were used to examine subtest relationships.

Reliability Estimates

Reliability refers to the consistency with which assessments and assessors measure behavioral attributes (Baumgartner, 1989). Reliability of the CMSP was examined through determining (a) test reliability (also referred to as internal consistency) with and without distraction scores as a factor and (b) interobserver reliability (also referred to as objectivity) (Baumgartner, 1989).

Reliability of the CMSP. Data for children (N = 297) from the 22 preschools were included in the reliability analysis. Reliability estimates were completed for (a) locomotor subscale, object control subscale, and total scores, (b) two and four trials, and (3) with and without distraction scores as a factor. Both Pearson correlations and intraclass correlation (ICC; one-way ANOVA) procedures were used. In estimating the effect of distraction scores on test reliability, "high" and "low" distraction categories were created using the median distraction score as the cut-off. Scores of 18 or higher were designated as high distraction, those less than 18 as low distraction. Separate analyses were run for high and low categories.

Interobserver reliability. To examine interobserver reliability, both testers scored performances on 50 children from 22 schools throughout the duration of the project. This represents approximately 11% of total participants (32 boys, 18 girls; mean age 4.3 ± .9 years). During reliability sessions, the testers were in close proximity; this was to ensure that they had similar perspectives for observing performance and recorded the scores concurrently and independently. Kappa statistics were used to provide estimates of interobserver reliability; estimates of Kappa were calculated for each process characteristic using data from the first trial only. Pearson correlations, the

TABLE 3
Concurrent Validity: Coefficients Among Motor Skill Assessments

	Locomotor	*Object Control*	*Total*
TGMD-2 vs. *CMSP-2*	.98	.97	.98
TGMD-2 vs. *CMSP-4*	.95	.94	.97
CMSP-2 vs. *CMSP-4*	.98	.97	.98

Note: TGMD-2 = original TGMD; CMSP-2 = CHAMPS protocol with two trials; CMSP-4 = CHAMPS protocol with four trials; $p < .0001$ for all values.

intraclass correlation (ICC) procedure, and the Estimates of Pearson correlation and ICC were completed for (a) locomotor skills, object control, and total test items, and (b) for 2 and 4 trials.

RESULTS

Concurrent and Construct Validity

To address concurrent validity, scores on the original TGMD-2, the CMSP with two trials, and the CMSP with four trials were compared using Pearson correlations. Coefficients ranged from .94–.98 for all comparisons (see Table 3). With regard to construct validity and the factor of "age differentiation," means and standard errors of measurement (SE) of 3-, 4-, and 5-year-old children are shown in Table 4, with and without control of distraction scores. There were significant differences among age groups for the locomotor subscale, the object control subscale, and total test scores for both four and two trials. Four-year-olds had significantly higher scores than 3-year-olds and 5-year-olds had significantly higher scores than 4-year-olds. Differences between 3- and 4-year-olds were greater than those between 4- and 5-year-olds.

CMSP Reliability

Reliability estimates for the CMSP are shown in Table 5. Estimates were high and ranged from $R = .88$ to .97. Reliability estimates when distraction scores were not accounted for were .90 and above and were slightly higher for four trials than for two. Reliability estimates based on distraction scores were similar for both high and low distraction categories; estimates in both categories were .88 or higher and, again, were slightly higher for four than for two trials.

Interobserver reliability. The range of Kappa coefficients for the first trial of each skill were: run (.80–1.00), broad jump (.48–.96), slide (.48–.88), gallop (.71–1.00), leap (.79–.84), hop (.77–.92), overarm throw (.50–1.00), underhand roll (.63–1.00), kick (.66–1.00), catch (.51–.92), strike (.32–.92), and dribble (.74–1.00) Due to missing data, Kappa coefficients for one process characteristic of the strike and one for the underhand roll could not be calculated. It is important to note that most skills had kappa coefficients above .6 for all process characteristics; four skills (jump, slide, catch, and strike) had one process characteristic with a kappa coefficient less than .6, and one skill (overarm throw) had two coefficients less than .6. Interrater reliabilities

TABLE 4
Age Differences Among CMSP Scores for Two and Four Trials

	CMSP-2 (Mean [SE]) Distraction Score Not Accounted For*			*CMSP-2 (Mean [SE])* Distraction Score Accounted For*		
	LM	*OC*	*Total*	*LM****	*OC*	*Total***
3 (n = 98)	31.8 (1.2)	37.6 (1.0)	69.5 (1.9)	31.9 (1.1)	37.7 (1.0)	69.6 (1.7)
4 (n = 148)	41.5 (1.0)	47.4 (0.8)	88.9 (1.6)	41.5 (1.0)	47.3 (0.8)	88.9 (1.4)
5 (n = 51)	46.7 (1.5)	54.7 (1.4)	101.6 (2.4)	46.4 (1.5)	54.4 (1.4)	101.0 (2.3)
	CMSP-4 (Mean [SE]) Distraction Score Not Accounted For*			*CMSP-4 (Mean [SE])* Distraction Score Accounted For*		
	LM	*OC*	*Total*	*LM*	*OC***	*Total***
3 (n = 98)	62.7 (2.4)	74.8 (1.9)	138.0 (3.8)	62.9 (2.3)	75.1 (1.9)	138.2 (3.6)
4 (n = 148)	82.2 (2.1)	94.9 (1.5)	177.2 (3.3)	82.2 (2.0)	94.8 (1.5)	177.1 (3.0)
5 (n = 51)	93.3 (3.0)	109.2 (2.6)	202.8 (4.8)	92.7 (3.0)	108.6 (2.6)	201.5 (4.7)

Note: CMSP-2 = CHAMPS protocol with two trials; CMSP-4 = CHAMPS protocol with four trials; LM = locomotor; OC = object control; Total = total score.

*All age groups different for all conditions $p < .0001$.

**Distraction score significant $p < .05$.

***Distraction score significant $p = .055$.

TABLE 5
Reliability of the CHAMPS Motor Skill Protocol (CMSP) for Two and Four Trials and Low and High Distraction Scores

	Without Accounting for Distraction Score		*Accounting for Distraction Score*			
	Two Trials	*Four Trials*	*Two Trials—High*	*Two Trials—Low*	*Four Trials—High*	*Four Trials—Low*
Locomotor	.92	.96	.92	.92	.96	.95
Object control	.90	.92	.91	.88	.95	.94
Total	.94	.97	.94	.94	.97	.97

Note: Scores of 18 or higher were designated as low distraction; those less than 18 were high distraction.

for two and four trials are given in Table 6. Results of ICC analyses indicated that interobserver reliabilities for four trials were high for the locomotor subscale (R = .99), object control subscale (R = .98), and total test (R = .94). Pearson correlations were also high and significant (p = .001 [locomotor subtest: r = .98; object control subtest: r = .97; total test: r = .90]). Correlations and ICCs for four trials were similar to or exactly the same as those for two trials.

The effect of "distraction score" on performances on the CMSP (four trials) was significant for both the object control subscale and total test scores ($p < .05$); distraction scores influenced the motor skills scores of children of all ages. For two trials, the effect of distraction

TABLE 6
Interobserver Reliability (Objectivity) for the CMSP

Test Components	*R*	*Raters Differ p-Value**	*Pearson Correlations*
Two trials (n = 50)			
Locomotor	.99	p = .66	.97**
Object control	.98	p = .29	.97**
Total	.94	p = .71	.89**
Four trials (n = 50)			
Locomotor	.99	p = .57	.98**
Object control	.98	p = .25	.97**
Total	.94	p = .77	.90**

*p-values for difference between testers.
**$p < .001$.

was significant for total test score only ($p < .05$). When distraction scores were controlled for, scores were slightly lower for 5-year-olds and slightly higher for 3-year-olds than when they were not (see Table 6). Overall the distribution of “distraction scores,” which ranged from 4 to 20, was skewed in the direction of “non-distracting” environmental conditions (mean score = 17.0, SD = 2.8, median = 18.0, minimum = 10, maximum = 20). The high mean score suggests that, in general, the environments in which testing took place were not distracting.

In terms of subtest correlations, another indicator of construct validity, relationships between subtest scores and total test score were high, positive, and in the appropriate direction. For four trials, the correlation of both object control and locomotor scores with total score was R = .88. The correlation between locomotor and object control subscale scores was R = .56. Correlations for four trials were similar to those for two trials (locomotor and object control subtest scores with total test score, R = .88; locomotor with object control subscale scores, R = .55).

DISCUSSION

Overall, data from this study indicate that the CMSP may be used as a reliable and valid method for assessment of preschool children’s motor skill performance in field-based settings. Specifically, test reliability, interobserver reliability, and validity coefficients were high and generally above .90. In terms of construct validity, evidence indicated that the CMSP differentiated among preschool age groups on locomotor, object control, and total test scores and, thus, provides a valid method for assessing age differences in gross motor skill performance in an important period of child growth and development. Subscale scores also correlated moderately with each other and were highly correlated with total test score, another indicator of construct validity. Concurrent validity was high and based on the positive correlation of the CMSP with the TGMD-2.

Although four trials were included in the protocol to provide assessors additional observations of a complex number of movement characteristics, data indicate that reliability and validity were similar for two and four trials. This suggests that comparable results may be obtained using two trials on the CMSP. Given the potential complexity associated with observing motor skill performances of large numbers of young children in field-based settings, consideration might be given to including four trials if testers are less experienced than the two CMSP testers, who had strong backgrounds in motor skill development, several years of experience working with young children, and extensive training prior to administering the protocol in the field setting. Lack of experience in the testing of motor skills of young children may be more likely for data collection staff involved in large-scale trials. In such cases, some preliminary training may be required. Another consideration may be the time required to administer two versus four trials; the time to administer four trials was approximately 45 min, while the estimated time to administer two trials is 25–30 min.

The number of movement skill characteristics to be observed did not appear to be an issue. This is a strength of the CMSP in that by formulating more precise, observable behavioral definitions of the movement skill characteristics to be observed, testers had clearer behavioral cues to assist them in "what to look for" in children's performances. These "concrete cues" may have aided test administrators in processing numerous movement skill characteristics and most likely contributed to the high test reliability, interrater reliability, and validity of the CMSP. These more precise behavioral definitions allowed for greater success in training testers to use the protocol reliably.

Environmental distraction scores had no affect on reliability of the test or on interobserver reliability (objectivity). This suggests that ongoing distractions in the testing environments did not, in this study, affect reliability of the assessment of children's performances. Data clearly were skewed toward a non-distracting environment (mean = 17); the range of scores also was limited (min = 10, max = 20, scale = 4–20). It is interesting that although reliability and objectivity estimates were not affected by "distraction scores," performance outcomes on both locomotor and total test scores were. These differences in performance outcomes were small but were present for children of all ages. The value of the distraction score is that it allows investigators to describe or rate and control for, at a simple level, the potential influence of the testing environment on measurement and performance outcomes. This is an advantage of the CMSP, since most protocols do not provide even minimal standardized guidelines for taking into effect environmental influences on assessment of performance outcomes.

A major strength of the CMSP is that it was used with a relatively large and diverse sample of preschool children from a variety of preschool circumstances and settings. A limitation of our study may be the method by which concurrent validity was established. The recommended approach is to administer the current test and the criterion test in two different time frames. In the present investigation, we collected data on the CMSP and the TGMD-2 within the same time frame. Nevertheless, coefficients were high and suggest that the CMSP is a valid measure of gross motor performance of preschool age children.

Overall, outcomes on the CMSP indicate that the protocol is appropriate for use in large-scale research to assess preschool children's gross motor skill performances and provides a viable alternative to existing instruments for use in epidemiological, population-based studies that involve large numbers of children. To gain a fuller picture of motor skill development of preschoolers, work is being initiated that will explore the use of developmental sequences in the protocol by examining a smaller subset of the preschool population. In addition, pending further replications and additional validation work, it is possible that the CMSP may also be a valuable assessment instrument for providing important information about children's gross motor skill needs, for making appropriate referrals for services to enhance children's motor skill development, and for recommending well-targeted and individualized motor skill development interventions where appropriate.

REFERENCES

Baumgartner, T. A. (1989). Norm-referenced measurement: Reliability. In M. J. Safrit & T. M. Wood (Eds.), *Measurement concepts in physical education and exercise science*. Champaign, IL: Human Kinetics.

Bishop, D. V., Ross, V. A., Daniels, M. S., & Bright, P. (1996). The measurement of hand preference: A validation study comparing three groups of right-handers. *British Journal of Psychology, 87* (Pt 2), 269–285.

Bricker, D., & Squires, J. (1999). *Ages and Stages Questionnaires: A parent-completed child-monitoring system*, (2nd ed.). Baltimore: Paul H. Brookes Publishing Co.

Burton, A., & Miller, D. (1998). *Movement skill assessment*. Champaign, IL: Human Kinetics.

Burton, A. W., & Rodgerson, R. W. (2001). New perspectives on the assessment of movement skills and motor abilities. *Adapted Physical Activity Quarterly, 18*, 347–365.

Bushnell, E. W., & Boudreau, J. P. (1993). Motor development and the mind: The potential role of motor abilities as a determinant of aspects of perceptual development. *Child Development, 64*, 1005–1021.

Clark, J. (1994). Motor development. In V. S.Ramachandran (Ed.), *Encyclopedia of human behavior* (pp. 245–255). San Diego: Academic Press.

Clark, J. (2005). From the beginning: A perspective on movement and mobility. *Quest, 57*, 37–45.

DeAgostini, M., & Dellatolas, G. (2001). Lateralities in normal children ages 3 to 8 and their role in cognitive performances. *Developmental Neuropsychology, 20*, 429–444.

Doty, A. K., McEwen, I. R., Parker, D., & Laskin, J. (1999). Effects of testing context on ball skill performance in 5-year-old children with and without developmental delay. *Physical Therapy, 79*, 818–826.

Fisher, A., Reilly, J. J., Kelly, L. A., Montgomery, C., Williamson, A., Paton, J. Y. et al. (2005). Fundamental movement skills and habitual physical activity in young children. *Medicine and Science in Sports and Exercise, 37*, 684–688.

Folio, M., & Fewell, R. (2000). *Motor development: Peabody Developmental Motor Scales (PDMS-2)*. Austin, TX: PRO-ED.

Gallahue, D., & Ozmun, J. (2006). *Understanding motor development: Infants, children, adolescents, and adults* (6th ed.). Boston: McGraw-Hill.

Gentile, A. (2000). Skill acquisition: Action, movement, and neuromotor processes. In J. H. Carr & R. B. Shepherd (Eds.), *Movement science: Foundations for physical therapy* (2nd ed.). Rockville, MD: Aspen.

Graf, C., Koch, B., Kretschmann-Kandel, E., Falkowski, G., Christ, H., Coburger, S. et al. (2004). Correlation between BMI, leisure habits and motor abilities in childhood (CHILT-project). *International Journal of Obesity and Related Metabolic Disorders, 28*, 22–26.

Henderson, S. E., & Sugden, D. A. (1992). *Movement assessment battery for children*. Sidcup, Kent: The Psychological Corporation, Ltd.

Leconte, P., & Fagard, J. (2004). Influence of object spatial location and task complexity on children's use of their preferred hand depending on their handedness consistency. *Developmental Psychobiology, 45*, 51–58.

Lyman, H. (1997). *Test scores and what they mean* (6th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Ogden, C. L., Carroll, M. D., Curtin, L. R., McDowell, M. A., Tabak, C. J., & Flegal, K. M. (2006). Prevalence of overweight and obesity in the United States, 1999–2004. *Journal of the American Medical Association, 295*, 1549–1555.

Piaget, J. (1963). *The origins of intelligence in children*. New York: Norton.

Sporns, O., & Edelman, G. M. (1993). Solving Bernstein's problem: A proposal for the development of coordinated movement by selection. *Child Development, 64*, 960–981.

Ulrich, D. (2000). *Test of gross motor development. Examiner's manual* (2nd ed.). Austin, Texas: PRO-ED.

Wiart, L., & Darrah, J. (2001). Review of four tests of gross motor development. *Developmental Medicine and Child Neurology, 43*, 279–285.

Wickstrom, J. (1983). *Fundamental motor patterns* (3rd ed.). Philadelphia: Lea and Febiger.

Williams, H. (1983). *Perceptual and motor development in young children*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Williams, H., & Monsma, E. (2006). Assessment of gross motor development in preschool children. In B. Barcken & R. Nagle (Eds.), *The psychoeducational assessment of preschool children*. Hillsdale, NJ: Lawrence Erlbaum.

Williams, P. T. (2001). Physical fitness and activity as separate heart disease risk factors: A meta-analysis. *Medicine and Science in Sports and Exercise*, *33*, 754–761.

Wilson, P. H. (2005). Practitioner review: Approaches to assessment and treatment of children with DCD: An evaluative review. *Journal of Child Psychology and Psychiatry*, *46*, 806–823.

Wood, T. M. (1989). The changing nature of norm-referenced validity. In M. J. Safrit & T. M. Wood (Eds.), *Measurement concepts in physical education and exercise science*. Champaign, IL: Human Kinetics.