Developing a General Outcome Measure of Growth in Movement for Infants and Toddlers

Charles R. Greenwood,
Gayle J. Luze,
Gabriel Cline,
Susan Kuntz, and
Carol Leitschuh
Juniper Gardens Children's Project, Early Childhood Research Institute for Measuring Growth and Development, University of Kansas

Proficiency in movement, an important outcome in early childhood, is necessary for physical, cognitive, and social–emotional development. The development of an experimental measure for assessing growth in movement in children ages birth to 3 years is described. Based on general outcome measurement (GOM) procedures (e.g., Deno, 1997), the measure was intended for the identification of children having difficulty acquiring movement skills and the evaluation of the effectiveness of interventions for these children. Results from its use with a sample of 29 infants and toddlers demonstrated the feasibility of the measure. The 6-minute GOM was found to be reliable in terms of interobserver agreement and odd–even and alternate forms indices. Adequate criterion validity was demonstrated through use of a standardized measure of motor abilities and caregiver ratings of children's movement skills. The GOM was sensitive to changes in key skill elements and growth in total movement rate with age. This Movement GOM holds great promise as a quick, inexpensive, and readily interpretable means of monitoring growth in movement for children from birth to 3 years.

Early childhood educators and early interventionists need information on the motor abilities and movement skills of young children. Practitioners need tools to identify children who may need early intervention in these areas. After children begin receiving interventions aimed at enhancing their abilities to move in their environments, practitioners also need tools to ensure that these interventions are appropriate and effective. Part C of the 1997 amendments to the Individuals with Disabilities Education Act (IDEA) stipulates that measures must be used to monitor children's progress toward goals outlined in the Individualized Family Service Plan (IFSP).

The terms *movement* and *motor* are sometimes used interchangeably, but they actually refer to different constructs (see, e.g., Burton & Miller, 1998). The term *movement* commonly refers to the observable behaviors involved in a change in posture or locomotion. The term *motor* commonly refers to the neuromuscular or other nonobservable internal processes or traits assumed to affect movement behavior. Because available assessment instruments differ with regard to the information they provide on movement skills and motor abilities, they also differ in the extent to which they inform the early intervention process. Because assessment of the movement of young children traditionally has been relegated to highly skilled occupational and physical therapists, early interventionists often face difficulties in obtaining information that is sensitive, timely, and relevant to the work they perform.

There are several critical reasons for assessing motor abilities and movement skills. One reason is to discriminate between very young children with delays in motor/movement skills and their peers without delays. A second reason is that delays in movement are often the first indicator that a child may have a more general developmental disability (Harries & McEwen, 1996). Delays in movement frequently appear more early in life than do delays in domains such as language or cognition. A third reason to focus on motor/movement assessment is that early problems in this domain often are related to developmental difficulties in other areas. Movement limitations may compromise children's abilities to communicate, interact, and explore their environments (Gallahue, 1989). This strong connection between movement and other

Address: Charles R. Greenwood, Juniper Gardens Children's Project, 650 Minnesota Ave., 2nd Floor, Kansas City, KS 66101; e-mail: greenwood@ku.edu

pro-ed

domains, especially cognitive development, was underscored by Piaget, who theorized that in the first 2 years of life, which he called the sensorimotor phase, motor actions in the environment provide the foundation for children's intelligence (Ramey, Breitmayer, & Goldman, 1984). For these reasons, we believe that early movement skills should be monitored to allow for early detection and intervention.

A number of approaches to assessing movement or motor skills in young children have been recommended in the early childhood literature. Harris and McEwen (1996) outlined three types of measures based on Kirshner and Guyatt's (1985) methodological framework for assessing health indices:

1. *discriminative indices,* which are used to distinguish between children with and without delays,
2. *predictive indices,* which are used to predict children's status (such as their ability to walk) at later ages, and
3. *evaluative measures,* which are used to assess change over time as a function of the intervention.

The first and third uses are the most important for practitioners.

Early interventionists face a number of challenges in obtaining this needed information, not the least of which are significant problems with existing measures. Most available measures are not usable by early intervention practitioners because they require the specialized expertise of occupational or physical therapists to administer and interpret results. The validity and utility of measures are highly linked to the theoretical framework on which they are based. For measures of motor abilities, frameworks typically posit the underlying neuromuscular system or the central nervous system. Too often, the evidence supporting the underlying frameworks for motor abilities is weak (Burton & Miller, 1998), and the treatment utility that this information provides practitioners is low. Alternatively, the *dynamic systems framework* posits movement development to be a function of the interaction of system variables, such as the person, the environment, and the task (e.g., Thelen & Smith, 1994). Dynamic system measures have more potential for informing early intervention because they tap the role played by environmental factors in motor development. One example of a measure based on the dynamic systems framework is the *Alberta Infant Motor Scale* (AIMS; Piper & Darrah, 1994). This measure is a 58-item criterion-referenced test of four motor abilities (prone, supine, sitting, and standing) and early movement milestones. Points for the individual items in the four scales are summed to form a total score that may be converted to a percentile. The AIMS is appropriate for use with children ages birth to 18 months of age, and it takes 20 to 30 minutes to administer. Information gathered through the AIMS regarding what a child can and cannot do is relevant to the design of movement interventions. A common practical limitation—the time required to administer measures such as the AIMS (ranging between 15 to 60 minutes depending on the measure)—reduces the utility for providing timely information to support ongoing treatment decisions. In addition, many assessments are deficient in basic psychometric features such as validity and reliability (Chandler, 1988). Most developmental assessments for movement or motor skills measure children's status at a single point in time; few directly measure growth or rate of progress over time.

Other measures—for example, *The Peabody Development Motor Scale–2* (PDMS-2; Folio & Fewell, 2000)—may not be appropriate because the environment needed to assess movement skills is not readily available to practitioners. For example, many childcare centers lack a set of stairs and handrails that meet the specifications required to administer items for measuring children's ability to climb stairs. In addition, measures may not sample the entire age range and the skill sets desired. For example, although the AIMS evaluates movement from birth through the age when children are walking independently, it does not include movement skills that need to be measured beyond infancy. Informal measures of movement, such as locally developed observations or ratings, may be practical and usable by practitioners but not generalizable to other children and other assessors because they lack standard administration procedures. Consequently, relatively few measures of motor and movement skills are appropriate for use by early intervention practitioners.

The General Outcome Measurement (GOM) approach has been recognized for its suitability for use by practitioners in identifying delays and monitoring progress (Deno, 1997). In this approach, an indicator of the general outcome is measured repeatedly and charted over time. Like pediatricians' height and weight charts, repeated measurement of suitable indicators provides pictures of individual growth over time relative to growth norms. Height and weight measures are quick, inexpensive, and easy to obtain reliably, making them useful to practitioners in the intervention decision-making process. Because growth in height and weight are sensitive indicators of the general outcome of "infant general health status," lack of expected growth rate signals concern, which sets the occasion for reflection on the problem, its causes, and possible solutions (McConnell, 2000). Lack of progress indicates the need for a more in-depth assessment to (a) determine in greater detail what the child can and cannot do and to (b) plan for an intervention or a change in an existing intervention. This approach has recently been used for assessing (a) the emerging literacy skills of young children learning to read (Kamin-

ski & Good, 1996) and (b) the expressive communication of children ages birth to 3 years (Greenwood, Luze, & Carta, 2002; Luze et al., 2001).

As part of a larger GOM development effort by the Early Childhood Research Institute on Measuring Growth and Development (ECRI-MGD; McConnell et al., 1996), this investigation sought to initially develop and validate a Movement GOM for use by practioners in assessing infants and toddlers from birth to 3 years. The framework used was one developed by Deno, Mirkin, and Chaing (1982), who described six criteria GOMs must meet:

1. GOMs should identify "authentic" child behaviors in natural settings. Authentic behaviors are particularly important for infants and toddlers because they are less able and willing to "perform" specific skills on demand. When asked to interact with unfamiliar people in unfamiliar situations, young children are even less likely to engage in desired behaviors. When children are given an opportunity to engage in typical behaviors in familiar environments, however, a much more accurate representation of skills can be obtained.
2. GOMs should assess key skill elements representative of an important child outcome.
3. GOM assessments should be standardized and replicable to ensure that the data from separate administrations are comparable.
4. GOMs must meet the requirements of technical adequacy, including interobserver agreement, internal consistency, and reliability of alternate forms, to provide accurate information that can be interpreted and used for decision making.
5. GOMs must be sensitive to growth over a short period of time so they can be used to evaluate intervention effectiveness.
6. GOMs should be efficient and economical, allowing practitioners to gather data that are usable for decision making without unduly adding to their workload.

This research investigated the feasibility of an experimental Movement GOM in terms of Deno's first five criteria. The specific research questions addressed focused on the measure's technical adequacy (i.e., reliability, validity) and sensitivity to growth over time.

Method

Participants

Children were recruited at two childcare centers serving infants and toddlers located in inner-city neighborhoods of metropolitan Kansas City. The centers served children of varied racial and socioeconomic backgrounds. Both centers were affiliated with neighboring high schools serving adolescent mothers. Any child in the center who was in the birth to 36-month age range was eligible to participate in the study. Each eligible child's parents received a packet of information that included an informed consent form and demographic questions. Any child whose parents returned a signed informed consent was included in the study. Thirty-nine informed consent forms (77%) granting parental permission were returned.

From this recruited sample, 34 children completed some aspects of measurement and 29 met a minimum criterion of three repeated GOM data points in the analysis sample. The mean age of these children at the start of the project was 15.3 months, with a range of 1 month to 34 months (SD = 9.6 months). Fifteen (52%) of these children were boys, and 14 (48%) were girls. For analytical purposes, the children were assigned to three age cohorts: birth to 12 months, 12 months to 24 months, and 24 months to 36 months, based on their age at the first measurement. The mean ages for these cohorts were 6.4 (n = 13), 18.4 (n = 10), and 29.5 (n = 6) months, respectively.

The racial breakdown of the sample was African American (84.7%), Hispanic/Latino (5.1%), European American (5.1%), and other or mixed races (5.1%). Five parents (17%) reported that their children were involved in Part C programs and had IFSPs. One of the children had Down syndrome; the other children had general developmental delays. Three children were at risk for a delay in movement as evidenced by PDMS-2 movement developmental quotient scores, which were at or below 0.78 (−1.5 standard deviations). One of these children was described by the teachers as having hydrocephaly.

Demographic information obtained from the parents revealed that annual family income levels ranged from $0 to more than $50,000. The modal income bracket was $10,000 to $17,000. Twenty-three percent of families had very low annual incomes (in the $0–$9,999 range). The mothers' highest educational level ranged from eighth grade to 6 years of post–high school education. The modal level of attained education was 11th grade, with 44% of mothers indicating that this was the highest level of completed schooling.

Measures

Criterion Movement Measures. Two criterion measures were used: One was a standardized test, the PDMS-2, and the other was a researcher-developed parental report measure, the *Caregiver Assessment of Movement Skills–Gross Motor* (CAMS-GM; Kuntz, 2001). These measures provided information from both examiners and parents/caregivers. The PDMS-2 is composed of six subscales: Reflexes, Stationary, Locomotion, Object Manip-

ulation, Grasping, and Visual–Motor Integration. It was administered by research staff members who had been specifically trained in its use. Administration required from 45 minutes to 60 minutes per child (Folio & Fewell, 2000). Only four of the six gross motor subscales were determined to be appropriate to the age range, however. Thus, the Stationary and Locomotion subscales were administered to all children, whereas the Reflexes subscale was used only with children 12 months or younger, and the Object Manipulation subscale (e.g., kicking, throwing) was used only with children 12 months and older. The reported overall reliability of the measure is 0.97; the reliability for the gross motor portion is 0.96. For the four subscales relevant to this investigation, reliability ranged from 0.84 to 0.96, with a mean of 0.92. In the analyses to be reported, raw, age-equivalent, and developmental quotient scores were used.

The CAMS-GM is a 40-item scale (39 objective items, 1 open-ended item) with items modeled after the *Bayley Scales of Infant Development–II* (Bayley, 1993), the *Denver Developmental Screening Test–II* (Frankenburg & Dodds, 1990), and the AIMS (see Piper & Darrah, 1993). We selected items that the parents would be able to recognize and judge whether their child could perform. The items were ordered developmentally. For example, the first item was "My child holds his/her head up to look around when lying on his/her stomach," and the last item was "My child kneels with both knees on the floor and returns to a standing position without using other objects to support getting up." Each item was marked by the parent as a skill that was *firm and well established* ("Does frequently/always," 2 points), was *just emerging* ("Does sometimes/occasionally," 1 point), was one the child *doesn't do* ("Does not do/Has never done," 0 points), or was one the parent had *no opportunity to observe* ("Don't know/not applicable," na). A single movement score was computed that reflected the mean rating of scored items. In order to remove bias contributed by lack of opportunities to observe a behavior, this calculation was adjusted by the number of items marked "NA" or (total points/[39 items − number marked "NA"]). This produced a movement mean score with a range from 0 to 2.0.

The Movement GOM. The Movement GOM was developed in two steps. In the first step, a general movement outcome was socially validated in a national survey of parents and professionals; in the second, the GOM was developed. Because an earlier study (Priest et al., 2001) had documented high, positive ratings for this measure by parents and professionals, the general outcome selected to guide development of the Movement GOM was as follows: "The child moves in a fluent and coordinated manner to play and participate in home, school, and community settings."

In the second step, a review of the movement and gross motor literature relative to infants and toddlers (Gallahue & Ozmun, 1995; Gifoyle, Grady, & Moore, 1981), as well as existing instruments (e.g., Miller & Roid, 1996) was conducted to guide development of the Movement GOM. Smith and Smith (1962; cited in Burton & Miller, 1998) proposed three categories of movement skills: postural, travel or locomotion, and manipulation. As Goldfield (1995) noted,

> The function of posture is to achieve stability, the state in which uncontrolled movements of the perception and action systems are minimized. By contrast, locomotion has two functions: (1) mobility, getting [an animal] from one place to another, and (2) exploration, discovering sources of needed environmental support and paths to them. (p. 250)

The function of manipulation is the control of objects through contact with, control over, and movement of them in specific contexts.

The literature further described the development of movement skills. At birth, infants have limited ability to control their movements: Almost all of their movements are reflexive and they have little ability to coordinate their arms, legs, and chest (Mowder, 1997). Between birth and 6 months of age, they become increasingly capable of moving independently to transition to a new posture within and across supine, side, and prone positions (Gilfoyle et al., 1981). They can typically lift their shoulders and chest and can roughly point to and grasp objects (Hannan, 1987).

Between 6 months and 18 months of age, most infants can sit up with support (Cratty, 1986). They are able to change positions by *pushing up, rolling,* and *extending their upper extremities.* The ability to *reach* and *corral* objects is emerging (Gallahue & Ozmun, 1995; Goldfield, 1995). Between 18 months and 24 months of age, movement becomes decidedly more complex and fluent as infants learn to *crawl, kneel, sit and stand without support,* and *stoop* (Burton & Miller, 1998; Gilfoyle et al., 1981). *Crawling forward or backward* on the belly typically advances to three- and four-point contact crawling positions (Gallahue & Ozmun, 1995). *Walking* first emerges as an upright unaided gait. Forward walking is soon accompanied by sideways and backward walking.

Between 24 months and 36 months of age, children are capable of hurried walk, can *step down* from low objects, and can *jump down* from an object and land on both feet. Most children achieve the ability to walk upstairs with help and then on their own. They also are able to walk downstairs unaided (Gallahue, 1989). Children are increasingly able to jump with both feet, engage in their first true *run, throw* objects like balls with force,

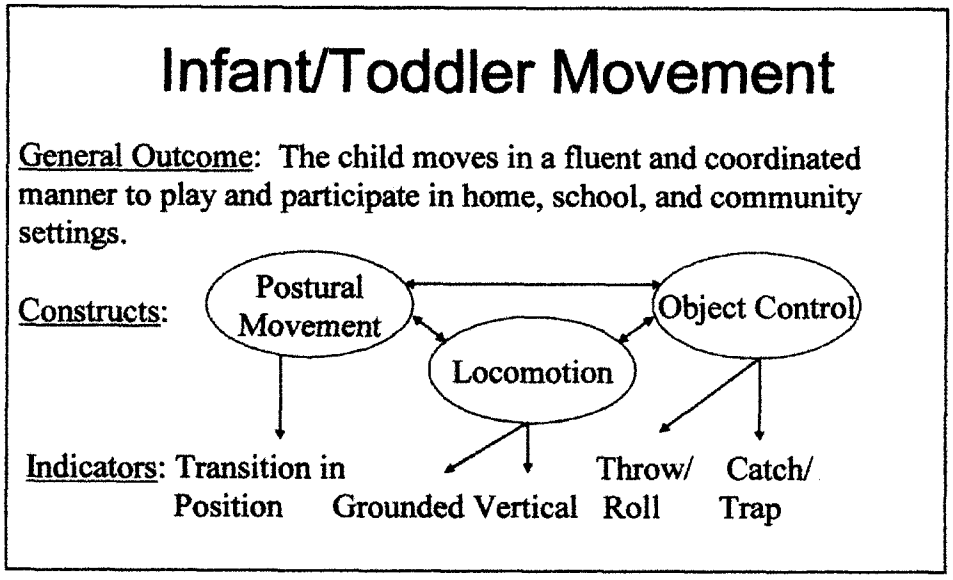

FIGURE 1. Conceptual framework linking the movement general outcome to theoretical constructs and indicators.

and *chase* the ball. They are also able to *trap or catch* an object coming at them, *kick* with a straight leg, and swing their arms to *strike* an object (Payne & Isaacs, 1991).

From this information, it was possible to identify key skill elements of movement that were linked conceptually and empirically to the general outcome and that were appropriate for measuring in children birth to 3 years of age (see Figure 1). As illustrated in Figure 1, the organization of movement skills for infants and toddlers included postural and transitional movement skills (e.g., rolling from back to stomach, weight shifting), locomotion skills such as grounded (e.g., belly sliding, crawling) and vertical locomotion (e.g., cruising, walking, running), and gross motor manipulation and object control skills (e.g., reaching, rolling/throwing, catching/trapping). Although one may presume a linear developmental path from postural to locomotion to object control in young children learning to move, it may be more appropriate to presume interactive, multidirectional paths whereby locomotion affects and is affected by postural movements that are affected by object control movements, and so forth.

A typical strategy for skill selection in the GOM approach is to sample for each domain or skill set of interest only *a few key skills* that are sensitive to increasing proficiency within the range of interest (e.g., Fuchs & Deno, 1991; Luze et al., 2001). In the area of movement, these selected skills included transitional movements, grounded locomotion, vertical locomotion, throwing/rolling, and catching/trapping (see Table 1). Because these same few skills are assessed on each and every measurement occasion, growth reflects a child's increasing proficiency in these skills from one occasion to the next. The GOM approach differs from the traditional criterion-referenced skills mastery approaches to assessment wherein *all skills* in a domain or task analysis hierarchy are tested on each measurement occasion (Fuchs & Deno, 1991).

In the pilot study, we tested several procedures for recording these key behaviors, including simple frequency count, percentage occurrence, and duration. Because the frequency count procedure produced greater numbers and captured more responding compared to the other methods, it appeared more sensitive and therefore better suited for further investigation. Next, we considered several testing situations in the form of settings and toy sets for evoking movement in young children. An initial effort was made to compare recording movement behavior frequencies in natural classroom situations versus structured, analogue play situations with familiar caregivers. Comparison of movement data of the same children measured in both situations indicated that the natural classroom situation was unsatisfactory because children typically moved much less there than in the analogue play situation. In addition, the frequency of movement behaviors varied considerably from day to day, depending on the availability of equipment, toys, and adults and peers within centers in the classroom. We therefore evaluated structured play situations composed of selected toy sets.

Toys were screened and selected based on their potential for engaging children's interest and evoking movement skills. Other important criteria for selecting toys included (a) safety, (b) availability in childcare settings, and (c) suitability for use for the entire age range of birth to 3 years. Toys that met the previous three criteria but were rejected too often evoked sitting and exploration

TABLE 1. Movement Key Skill Element Definitions

Key skill element	Definition	Example behaviors
Transitional movement	Transitional movements are motions used by a child to achieve a new position within a posture or to achieve a new posture. This can include movement within a stable posture (changing the primary weight-bearing surface) or moving from one distinct posture (Lying supine or prone, sitting, kneeling, stooping, standing) to another. An episode begins when a child begins moving from a stable position to a new position. The episode ends when the child has regained a stable position or begins locomotion.	Rolling to stomach from back, rolling to back from stomach, moving in and out of sitting position, standing up, kneeling down resting on knees
Grounded locomotion	Locomotion involves movements that transport the body forward, backward, sideways, or upward from one point in space to another. Grounded locomotion is movement horizontal to the ground and does not use upright postures when moving.	Moving on belly from one location to the next, either forward, backward, or sideways (pivot in prone); thrusting arms forward and then subsequently flexing them in a movement that results in a slight forward or backward movement
Vertical locomotion	Vertical locomotion is movement in an upright position that moves the child forward, backward, or sideways.	Cruising is walking while holding onto furniture for support; walking involves alternating feet where one foot is always on the floor
Throwing/rolling	Throwing is propelling an object through the air. Rolling is pushing a circular object so that it rolls away from the child's body.	Throwing an object using an overarm, underarm, or sidearm throw; rolling an object toward a person
Catching/trapping	Catching is bringing an airborne object under control using hands and arms. Trapping is stopping a moving object (moving through the air or rolling on the ground) with the hands, arms, legs, or body.	Catching an object with one's hands or arms or trapping it against the body

and less gross movement. Examples include Ball Party™, motorized walking toys, dinosaur pull toys, ramps, and cars.

The final list of toys acquired from Constructive Play Things™ and shown to evoke the most movement behaviors included blocks/balls (BB), the window house (WH), and the shopping cart (SC), along with enhancements such as a rattle, a squish toy, and squeaky food toys and blocks that popped together (see Figure 2). The BB toy set was composed of various sizes and shapes of foam rubber blocks along with balls of differing sizes and textures. For the youngest children, a rattle was used for grasping and throwing. A favorite activity with the blocks and balls was to build a tower with the blocks and rattle and then knock it over with the balls. The WH was a large foam vinyl square made by Environmental, Inc. It opened on both ends so that the children could crawl through it. The WH included two see-through mesh-covered windows and one open window. In addition to crawling through, children could stand up in the house and look out the windows. Also included with the WH were some balls and a squish toy for the children to play with. The balls encouraged a lot of the movement in our initial work with the WH. The SC was a plastic, child-height, multicolored cart on wheels. Plastic food, hand-held musical instruments, and toys that popped together were added to give the children something to "shop for" and something that could be picked up, put into, and taken out of the cart.

The BB, WH, and SC toy sets—along with the movement recording protocol—were pilot tested further with 12 children stratified by age: under 1 year (n = 4), between 1 year and 2 years (n = 4), and between 2 years and 3 years (n = 4). Movement assessors recorded the frequencies of occurrence of each movement element: transitional movement, grounded locomotion, vertical locomotion, throwing/rolling, and catching/trapping. The frequencies of individual children's movements using these toys were assessed during a 6-minute play session. A total movement score, the sum of the individual movement frequencies, was calculated for each child. These scores were converted to rate (responses/min) by dividing by time observed (i.e., 6 minutes).

The children's observations generally took place with a familiar adult as a play partner in an available room that contained the toys and was removed from the distraction of the classroom. Previous experience at the centers indicated that location of assessment did not have a

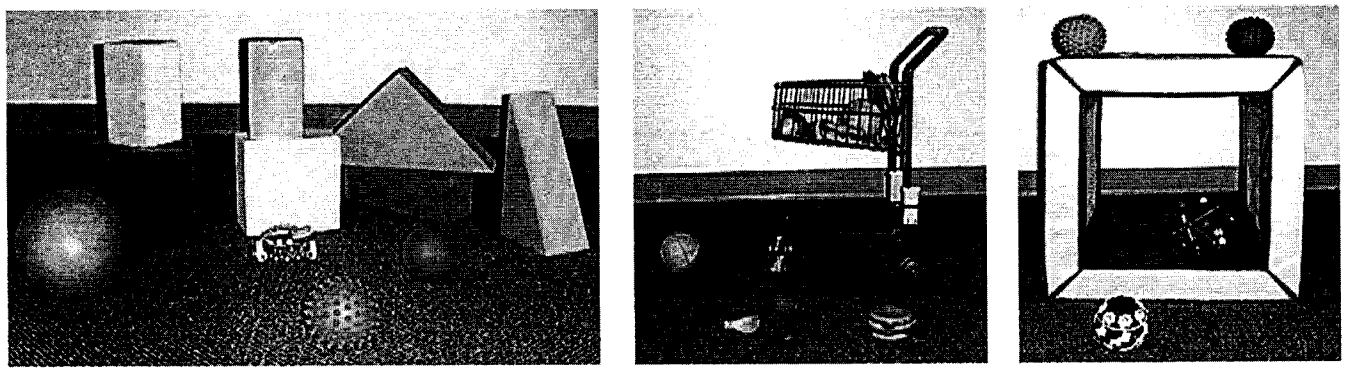

FIGURE 2. The three toy sets used as alternate forms. The blocks/balls toy set (left panel) was composed of foam blocks of various sizes and shapes and balls of varying sizes and textures. The shopping cart toy set (middle panel) was a multicolored, child-height plastic shopping cart on wheels. Also included were plastic food items, hand-held musical instruments, and pop-together toys as "to shop for items." The window house toy set (right panel) was a large, vinyl plastic square open at both ends with see-through mesh-covered windows that the children could crawl through and stand in. It came with balls of varying sizes and textures.

significant effect on the children's responses. In one childcare center, a general meeting room upstairs from the children's regular classroom setting served this purpose; in the other center, an infant sleeping room was used. The research staff set up the testing session with the toys and a videocamera for recording the session. They then would go and "warm up" the children in their familiar environment before bringing them to the testing session. During the GOM, an adult play partner, typically a research staff member, engaged the child in a play session lasting 6 minutes. During this time the play partner followed the child's lead while a second member of the research staff videotaped the child's behavior.

In fewer than 10 instances, when a child was wary of leaving the familiar environment without an adult caregiver, the adult caregiver sat on the perimeter of the testing situation where the child could still see him or her. In several additional cases, when children were hesitant to play with project staff members, their parent or caregiver acted as their play partner during the GOM assessment.

Design and Procedures

The general design was a longitudinal study with pretests and posttests of criterion measures separated by nine sequential GOM measures. Each of the GOM measures was separated by 3 weeks, resulting in a total study period of 8 months. After orienting and recruiting of prospective childcare center staff members had been completed, informed parental consent was obtained. Unfortunately, most children missed the second GOM wave of data because both centers were closed due to snow and winter break. Sixty-six percent of the 29 children were measured eight times, 2% 7 times, 17% six times, 3.5% four times, and 7% three times. Prior to and after these measures, the PDMS-2 and CAMS-GM measures were administered, and data were gathered. Twenty-two children had complete PDMS-2 data for both Time 1 and Time 2, but only 13 had complete data on the CAMS-GM.

The three current alternate forms (SC, BB, and WH) were varied systematically across sessions for each child in order to represent the child's movement performance with these different situations and thus provide a means of assessing the reliability of alternate forms. The order was counterbalanced across centers and repeated across subsequent measurement waves. The order at Center 1 was WH, SC, and BB, whereas at Center 2 it was SC, BB, and WH.

Each GOM assessment was videotaped. The tapes were subsequently coded using the key skill element definitions (see Table 1). These records constituted the primary GOM data collection and also assessments of interobserver agreement. After recording, frequency scores were tallied and entered into a computer database (i.e., Visual d-BASE) for subsequent conversion to rate per minute scores and for charting and analyses in other statistical software (i.e., MS-EXCEL, SPSS). Parents received gift certificates in the amount of $10 after completing the pretest parent measures; another $10 gift certificate was provided after the parent posttest measure was completed.

Staff and Training. Project staff included research assistants and staff assistants, all of whom had previous experience teaching young children. In addition, each of the staff members had a history of recent work in the childcare center in which they were conducting assessments, and each was familiar to the children. The assessors received training in the purpose of each experimental measure, the specific administration steps for the PDMS-2, and the GOM. To certify as trained, each staff member was required to obtain three 90% or higher interrater

agreement scores prior to collecting data for the study. The CAMS-GM was prepared and distributed at the childcare center to the parents.

GOM Interobserver Agreement. Throughout the study, interobserver agreement was assessed on a randomly selected 78 (38%) of 206 GOM assessments. Agreement was checked by a second person who independently recorded the same videotaped session simultaneously with the primary observer. The two observers' frequency counts for each key element and their total communication scores were compared for agreement within each agreement check. Because end-users were anticipated to be early childhood practitioners, an easy to calculate but liberal agreement calculation—the frequency ratio or marginal agreement statistic (Hartman, 1977)—was used. The frequency ratio was computed by dividing the smaller of the two scores into the larger scale and multiplying by 100. The mean frequency ratios were 93% (SD = 0.07) for total movement, 93% for transitional movement, 88% for grounded locomotion, 89% for vertical locomotion, 89% for throwing/rolling, and 86% for catching/trapping. Separate analyses of these data examined observer agreement regarding the occurrence and nonoccurrence of these behaviors. The nonoccurrence analyses produced consistently high levels of agreement above 0.85 on all indicators. With the exception of catching/trapping (0.60), the occurrence analyses were similarly high. The lower level of occurrence agreement for catching/trapping appeared to be due to the infrequent occurrence of these behaviors in general. For example, in only 28 of the 78 agreement checks did at least one observer record an occurrence of catching/trapping. An additional occasional problem was poor picture quality on the videotape.

To provide a conservative analysis of GOM score reliability, Pearson product-moment correlations were used to calculate the covariation between observers' frequency scores over the 78 agreement checks. These values were .98 for total movement, .97 for transitional movement, .96 for grounded locomotion, .95 for vertical locomotion, .95 for throwing/rolling, and .92 for catching/trapping. A final analysis compared the observers' mean estimates for each of these behaviors. Observers' mean values were highly similar and not significantly different in all cases, with the exception of catching/trapping: M = 0.54 responses per minute versus M = 0.65 responses per minute, $t(77) = -2.001$, $p = .05$.

Statistical Analysis. Simple descriptive statistics (means and standard deviations) were used to represent the magnitude of the various dependent measures and individual variations in the mean and over time. Pearson product-moment correlations, dependent t tests, and ANOVAs were used to test indices of reliability as well as validity. Hierarchical linear modeling (HLM2; Bryk & Raudenbusch, 1992; Bryk, Raudenbusch, Cheong, & Congdon, 2000) was used to examine individual and group growth parameters over time. Level 1 HLM analyses were used to compute slope (rate of growth) and intercept parameter values for individuals and groups. Level 2 HLM analyses were used to model growth as a function of age cohort (Cohort 1 = 0–12 months, Cohort 2 = 13–24 months, and Cohort 3 = 25–36 months), gender, childcare center, and disability status (IFSP: yes or no).

When it was suspected that a growth curve was not linear, the likelihood ratio test was used to assess the need for including an "acceleration" parameter in the growth model at Level 1 (Bryk & Raudenbusch, 1992). The likelihood ratio test is used to compare the deviance statistics of two nested models, with one model containing the parameter and the other not. Growth models with smaller deviance statistics better fit the actual data (Bryk & Raudenbusch, 1992).

Examining individual patterns of growth over time is a particularly appropriate application of HLM (Bryk & Raudenbusch, 1992; Hatton, Bailey, Burchinal, & Ferrell, 1997; Luze et al., 2001). HLM produces estimates equivalent to multivariate repeated measures (MRM) methods with several advantages: (a) it explicitly represents individual growth; (b) it has generally more flexible data requirements because repeated measurements are considered nested within the person; and (c) compared to MRM, it provides a more flexible specification of the covariance structure among repeated observations supporting hypothesis testing (Bryk & Raudenbusch, 1992). HLM also is flexible in accommodating missing data.

A unique advantage of HLM analysis is the ability to compute a mean intercept at a single point in time or test for mean differences between groups at a point in time (Bryk & Raudenbusch, 1992). Termed *centering*, calculating the intercept at a specific point in time produces a unique value for the intercept; however, the value of the linear slope is unaffected. Unless otherwise indicated, the intercept means in this study were centered at the fifth measurement occasion (the middle of the study) or at 18 months of age (the midpoint in the birth–36-month chronological age range).

RESULTS

Movement GOM Reliability

Odd–Even Reliability. Odd versus even reliability was calculated by consolidating data from each child's odd and even measurement occasions to form mean scores and correlating these point estimates across children. The odd–even Pearson product-moment correlation was $r(29) = 0.88$, $p < .001$, and the mean point estimates for

each were not significantly different: 8.8 versus 8.2 movements per minute, $t(28) = 1.46$, $p = .154$.

Alternate Forms Reliability. Alternate forms reliability was calculated by consolidating each child's observation scores within a common toy form. These forms were BB, WH, and SC. Pearson product-moment correlations between total movement estimates were computed. These correlations were relatively large and positive at 0.85 (BB vs. WH), 0.84 (WH vs. SC), and 0.91(BB vs. SC), and each correlation was statistically significant ($p < .001$). Total movement mean estimates were 9.2 versus 9.4 (BB vs. WH), 10.1 versus 6.8 (WH vs. SC), and 9.9 versus 7.0 (BB vs. SC). Paired t tests indicated no differences in total movement for BB when compared to WH; however, SC produced estimates that were on the order of three responses per minute lower than that of either BB, $t(24) = 4.73$, $p < .001$, or WH, $t(25) = 5.22$, $p < .001$.

Movement GOM Criterion Validity

PDMS-2/CAMS-GM Predictive and Concurrent Validity. The Time 1 and Time 2 means and standard deviations for the PDMS-2 and CAMS-GM measures are provided in Table 2. As can be seen, all indices increased 45 weeks later at Time 2, and these increases were statistically significant for all but the PDMS-2 developmental quotient, a relativistic metric rather than an isomorphic one. Several analyses examined relationships within the PDMS-2 and CAMS-GM report measures over time and between these criterion measures of movement at the same time points (see Table 3). On the PDMS-2, children's scores (stationary and locomotion) were significantly correlated from Time 1 to Time 2, reflecting stability in the children's relative rank order. This was true even when the children had grown in movement skills over time. For example, the Time 1 versus Time 2 correlation of the locomotion scale was $r = 0.96$, and the mean locomotion scale gain score was 17.4 raw score points, the raw score difference between the first and last assessments (80.0 minus 97.4). The stationary scale was similarly correlated across occasions (see Table 3). Similarly, the Time 1 versus Time 2 correlation of the caregiver report was significant at $r = 0.94$, and the raw score mean gain was 0.14, or 1.25 minus 1.39 (Time 1 and Time 2, respectively).

TABLE 2. PDMS-2 and CAMS-GM Descriptive and Inferential Statistics

		Time 1		Time 2					
Measure	Indicator	*M*	*SD*	*M*	*SD*	Gain	*F*	*df*	*p*
PDMS-2	Stationary[a]	21.18	17.17	27.23	18.86	6.05	16.19	1, 21	0.001
	Locomotion[a]	17.95	12.34	23.37	14.23	5.32	38.80	1, 21	0.0001
	Total[b]	102.55	18.48	105.41	18.83	2.86	2.36	1, 21	0.139
CAMS-GM	Rating[c]	1.25	0.65	1.39	0.58	0.14	4.74	1, 12	0.05

Note. PDMS-2 = *Peabody Developmental Motor Scales–2* (Folio & Fewell, 2000); CAMS-GM = *Caregiver Assessment of Movement Skills–Gross Motor* (Kuntz, 2001); GOM = general outcome measure.
[a]Age equivalent scores. [b]Developmental quotient. [c]Sample mean scores.

TABLE 3. Intercorrelations Among PDMS-2, CAMS-GM, and GOM Movement Measures

	Instrument	Indicator/occasion	1	2	3	4	5	6	7
1	PDMS-2	Stationary 1	1.00						
2	PDMS-2	Locomotion 1	0.87*	1.00					
3	PDMS-2	Stationary 2	0.90*	0.90*	1.00				
4	PDMS-2	Locomotion 2	0.93*	0.96*	0.93*	1.00			
5	CAMS-GM	Mean Rating 1	0.71*	0.78*	0.68*	0.76*	1.00		
6	CAMS-GM	Mean Rating 2	0.83*	0.76*	0.72*	0.78*	0.95*	1.00	
7	GOM movement rate	Intercept	0.80*	0.90*	0.77*	0.86*	0.85*	0.87*	1.00

Note. PDMS-2 = *Peabody Developmental Motor Scales–2* (Folio & Fewell, 2000); CAMS-GM = *Caregiver Assessment of Movement Skills–Gross Motor* (Kuntz, 2001); GOM = general outcome measure.
*$p < .01$.

The PDMS-2 and CAMS-GM measures also converged at both Time 1 and Time 2. These Pearson product-moment correlations ranged from 0.68 to 0.83 between measures, depending on the measurement occasion. They therefore each shared substantial common variance in describing children's movement.

Criterion Validity of GOM. Table 3 summarizes the extent to which the GOM total movement rate correlated with the PDMS-2 and CAMS-GM measures. These cross-measure correlations were positive, large, and statistically significant for the total movement intercept, r = 0.90 at Time 1 and r = 0.86 at Time 2, in relationship to the PDMS-2 locomotion scale. Correlations to the stationary scale scores on the PDMS-2 also were high and positive, r = 0.80 at Time 1 and r = 0.77 at Time 2. The correlations between the intercept of the GOM total movement rate and the CAMS-GM also were positive, large, and statistically significant, r = 0.85 at Time 1 and r = 0.87 at Time 2.

Sensitivity to Growth Over Time

Changes in Key Skill Elements by Age Cohort. Central to the construct validity of a GOM is the hypothesis that as children age, more sophisticated developmental skill elements will emerge and grow differentially. In fact, we expect new skills to emerge, existing skills to become more fluent, and old skills to attenuate or be replaced by other, more advanced skills. As seen in Figure 3 for the three age cohorts in the first year (upper panel), second year (mid panel), and third year (lower panel) of life, respectively, this was generally the case. The major movement activities for children between 3 months and 12 months (Cohort 1) were emergence of and growth in transitional and vertical locomotion, followed in order by grounded locomotion, throwing and rolling, and catching and trapping. For children between 13 months and 24 months (Cohort 2), these skills all occurred more often and continued to grow (transitional movement; throwing/rolling) or level off (vertical and grounded locomotion) with catching/trapping still emerging. For children between 25 months and 36 months (Cohort 3), these skills again occurred more frequently, but trends over time were declining (vertical locomotion) or leveling off (transitional movement; grounded locomotion). Throwing/rolling and catching/trapping continued to grow.

Growth in Total Movement Rate by Age. Central to a composite GOM movement indicator is growth over time. The HLM Level 1 analysis indicated significant mean intercept and slope. The mean intercept centered at the fifth measurement occasion was 8.44 total movements per minute, SE = 0.91, $t(28)$ = 9.23, p = .0001. The mean slope was 0.38 total movements per occasion, SE = 0.10, $t(28)$ = 3.95, p = .001. When translated to real time, the average number of movements across all children grew by 0.51 movements per month. The HLM Level 2 test of the effects of age cohort was significant for both mean intercept, $t(186)$ = 6.16, p = .0001, and mean slope, $t(25)$ = −3.04, p = .006. The mean intercepts were 5.26, 9.95, and 14.65 movements per minute for age Cohorts 1, 2, and 3, respectively. Similar values for the mean slope were 0.65, 0.35, and 0.05 for each cohort, respectively. As children aged, the rate of growth declined, suggesting an accelerating then decelerating curve (quadratic) rather than a linear growth curve over the entire age range (see Figure 4).

Given these findings, a quadratic term was added to the growth model describing the complete age range. The model containing the quadratic term produced a smaller deviance statistic, 1108.25 (with 10 estimated parameters) versus 1126.13 (with 6 estimated parameters), and the likelihood ratio test (Bryk & Raudenbusch, 1992) was significant, $\chi^2(4, N = 29) = 17.88$, p = .002, indicating that the fit to the data was better with the quadratic term included in the model. The resulting growth curve parameters in the quadratic model for movements per minute were as follows: 7.89, $t(28)$ = 12.085, p = .001, for intercept; 0.559, $t(28)$ = 15.16, p = .001, for slope; and −0.013, $t(28)$ = −3.262, p = .003, for acceleration. This growth curve is displayed in Figure 5, with the intercept centered at 18 months. Also displayed in Figure 5 are the growth curves defining the negative and positive 1.0 standard deviation interval around the mean growth trajectory. In addition, the actual trajectories of the three children identified on the PDMS-2 as at risk for a developmental delay in movement are shown. With minor exception, the GOM trajectories of these children who were low performing on the PDMS-2 typically fell below the mean GOM total movement rate trajectory, and in the case of Child 123 and Child 135, they fell consistently below the −1.0 SD curve on total movement rate.

Differences in Total Movement by Gender, Center, and IFSP Status. Additional Level 2 tests for gender, childcare center, and disability as indicated by the IFSP indicated no significant differences in growth parameters.

Discussion

This research investigated the feasibility of an experimental Movement GOM. Its technical adequacy and sensitivity to growth over time were investigated. Of particular importance was the fact that the 6-minute GOM proved to be generally reliable in terms of (a) interobserver agreement and (b) odd–even and alternate forms indices. The experimental GOM also attained adequate criterion validity in relationship to a standardized measure of motor abilities (PDMS-2 locomotion and stability scales) and to caregiver ratings of children's movement skills (CAMS-

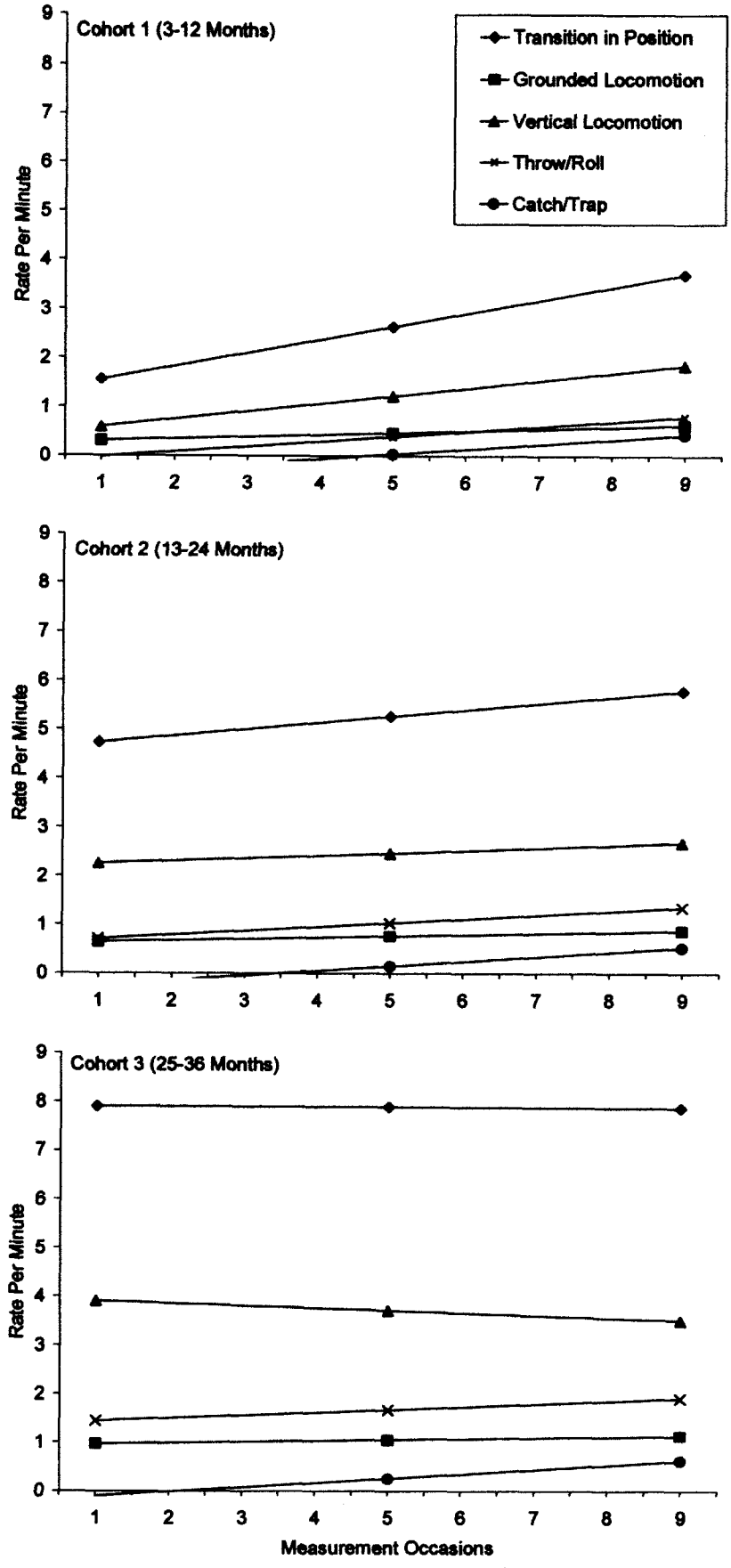

FIGURE 3. Growth in the rate of key movement skill elements over measurement occasions (separated by 3 weeks) by age cohort.

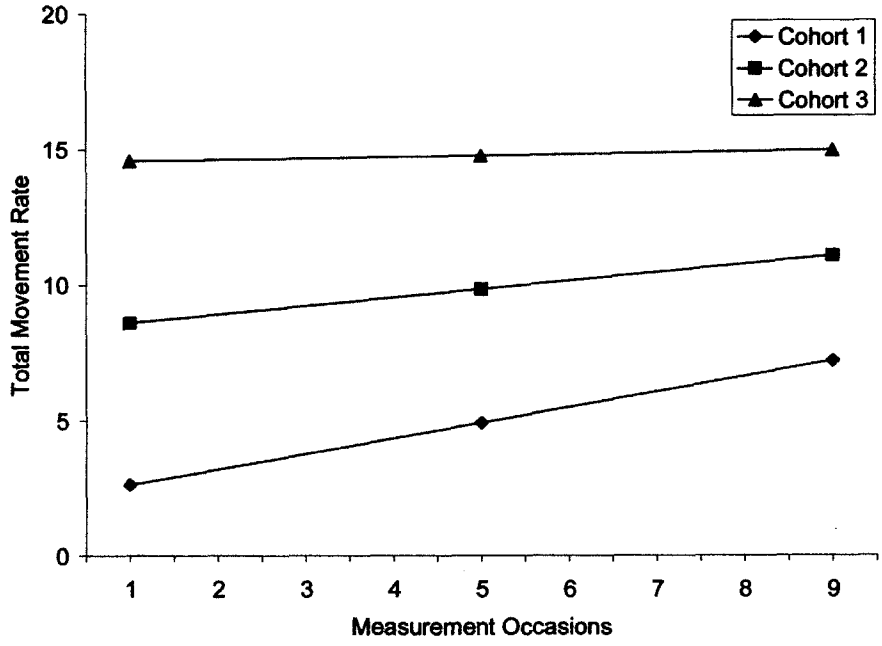

FIGURE 4. Growth in total movement rate over measurement occasions (separated by 3 weeks) by age cohort.

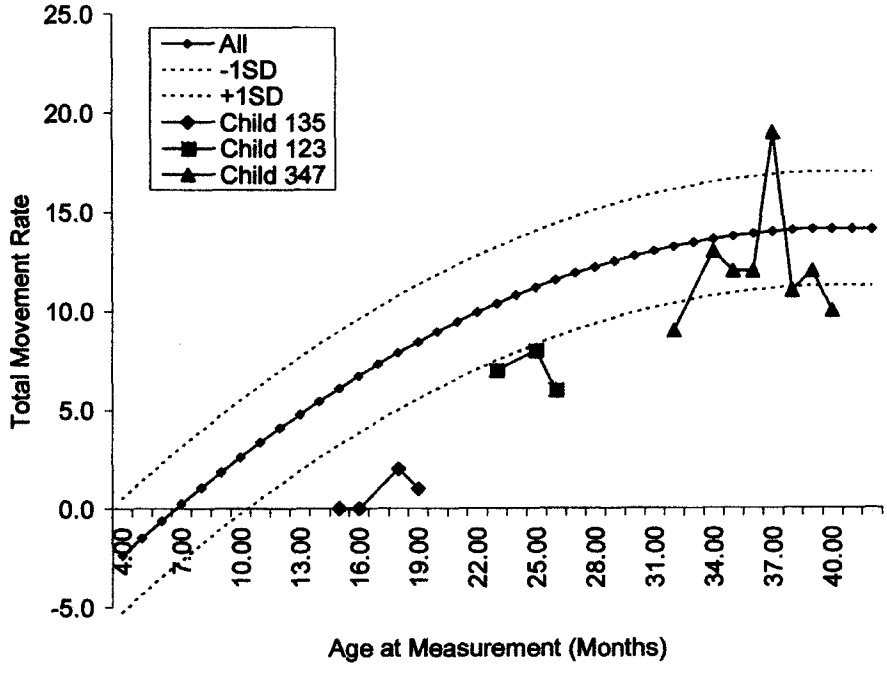

FIGURE 5. Growth in the rate of total movement by age at measurement (in months) for all children, by + 1 *SD* and − 1 *SD* of the mean intercept, and for children at risk for movement delays per the *Peabody Developmental Motor Scale–2* (Folio & Fewell, 2000).

GM), indicating that the GOM measured movement in the context of a multimethod, multi-informant validation framework. A multimethod framework is valued because of the convergence of several measures rather than the use of a single measure (e.g., Lewin, Hops, Davis, & Dishion, 1994). In addition, for children who had a discrepancy of the −1.5 *SD* on the PDMS-2 and who had IFSPs, performance on the Movement GOM also indicated normatively lower mean levels and rates of growth than expected based on the performance of the entire sample of children (see Figure 5). The GOM was sensitive to change in key skill elements and to growth in total movement rate with age. Supporting the GOM's construct validity is the fact that older children produced more frequent rates of movement per key skill element and total rate of movement than did younger children. At least some key skill elements increased across time within each age cohort.

One concern was the high level of missing parent data on the CAMS-GM in the second wave of measurement. This occurred at the end of the program semester in advance of summer vacation. Parents did not return the questionnaire despite repeated contact efforts. The net effect of this problem, however, was limited to correlations of measures only at the end of the study at Wave 2. A more complete set of data was available in the first wave, and variations between Wave 1 and Wave 2 related to this problem appeared to be minimal.

Results argue strongly in favor of the technical adequacy of the Movement GOM as a measure of movement development. More important is its potential contribution to the assessment of the movement skills of infants and toddlers compared to currently existing measures (e.g., McConnell, 2000). With its focus on movement skills rather than motor abilities, the GOM provides information relevant to the use by early intervention practitioners of movement-promoting activities conducted in the context of families' routines or activities in early education centers. Because the GOM provides information on progress and rate of growth in movement, it has the potential for aiding in the creation and validation of new interventions that could lead to improved movement development. Compared to the motor abilities and criterion skills mastery assessment approaches, the focus in the GOM is on proficiency in a few strategic skills selected to indicate progress toward attaining the general outcome. Whereas the other approaches yield comprehensive profiles of milestones (motor abilities) or of skills mastered at some level of quality (criterion-referenced) on each measurement occasion, this is not the intention of the GOM. Instead, growth in performance on the indicator is produced at each measurement. Detailed information about skills and skills quality is needed in the GOM approach in several situations: when progress is lacking, when designing an intervention plan, when exploring solutions, or when changing an intervention plan that is not working (McConnell et al., 1998). In such cases, information from measures like the AIMS that can identify what skills need to be learned and taught is highly relevant; however, progress resulting from associated interventions should continue to be monitored by the GOM (Fuchs & Deno, 1991).

The Movement GOM appears to have a number of advantages and few disadvantages in comparison to other measures of movement. One obvious advantage is the shorter 6-minute administration time compared to other established movement measures such as the AIMS, which takes 20 minutes to 30 minutes. The Movement GOM covers 3 years of development, whereas the AIMS covers 18 months or until children walk independently. Another advantage is the short time allowed between administrations (3 weeks to 4 weeks for the GOM vs. 6 months for the PDMS-2). The GOM also provides a rate of growth metric (slope) rather than a gain in total score or change in percentile rank, as in the AIMS. Finally, the Movement GOM requires a lower level of training and is easier to interpret as compared to the AIMS.

The GOM and AIMS have several problems in common. Both suffer from limited normative groups. The AIMS was normed in Canada and may not be representative of children in the United States. The GOM currently lacks a normative sample, and the small sample used in this study was primarily African American. A Movement GOM normative sample awaits future research. Both measures appear to be influenced by ceiling effects. In the AIMS, the total score is a function of mastery of the total number of items and is therefore a fixed ceiling. Current findings indicate that the Movement GOM is most sensitive to the growth in total movement of the children in Cohort 1 (slope = 0.65), declining by about half for Cohort 2 children (slope = 0.35) and to nearly zero (slope = 0.05) for Cohort 3 children. When plotted over all children by age at measurement, the net effect was a curvilinear growth trajectory that accelerated for the younger children while slowing for the older children (see Figure 4). This is not the only case reported in the literature of a GOM declining in sensitivity with age and increased level of fluency. For example, Fuchs, Fuchs, and Hamlett (1993) reported that oral reading rate, a widely used elementary grade–level reading GOM, declined in sensitivity over increasing grade levels.

It was not immediately obvious why the oldest cohort had the slowest rate of growth. Because the GOM is not a skills mastery measure, it is not prone to ceiling effects in the same way as the AIMS. A plausible explanation was that children who were highly fluent in these skills reached a ceiling effect as each movement became of longer and longer duration, which reduced frequency. Children who are fluent in vertical locomotion, for example, might continue walking without stopping for sev-

eral seconds as compared to a child who is just beginning to move in an upright position. In order to measure growth in movement more sensitively for these more highly skilled and fluent "movers," future studies may need to focus on key skill elements in duration of movement. The interrelationship between frequency and duration metrics requires further research, with the caveat that the final product must remain an observational procedure simple enough for use by childcare practitioners and early interventionists.

With respect to improving this GOM's sensitivity for children in the older cohorts, more work needs to be done. First, although the three toy forms produced acceptable high levels of intercorrelation, one of the toy forms (SC), evoked significantly lower rates of movement behavior in comparison to the other two forms. Its use in combination with the other two forms thus added to the variability in each child's growth trajectory as reflected in standard error terms. Small ratios between growth parameters (e.g., mean intercept and slope) and their standard errors of estimate are desired in GOMs (Fuchs & Fuchs, 1999). Additional research is needed that replaces the SC context with a truly equivalent form. Second, in terms of key skill elements, growth in rolling/throwing and growth in trapping/catching were least evident in each cohort; in fact, they rarely reached above one response per minute, even in the oldest children in Cohort 3. Future work should focus on the toy and play features of the toy forms and attempt to achieve situations that better evoke a display of these skills.

Conclusions

This research advanced both the case for the feasibility of developing and validating movement GOMs for infants and toddlers and the broader case for wider use of GOM measurement in the field of early intervention. Several criteria for validating a GOM were addressed (Deno et al., 1982). First, authentic movement behaviors were assessed in a childcare setting. Second, key skill elements of movement related to the general outcome were measured. Third, standardized procedures were developed and used repeatedly over time. Fourth, the Movement GOM was shown to be largely technically adequate. Fifth, the Movement GOM was shown to be partially sensitive to growth over time.

Future work must replicate and extend the present findings to larger samples. Future research also must address the feasibility and utility of the Movement GOM when used by practitioners. In particular, treatment validity, or the extent to which it is sensitive to intervention effects over time, will be of interest. In the present research, children's data were coded from videotapes of 6-minute sessions. In order for the Movement GOM to be used effectively in childcare settings, such data will necessarily need to be collected by practitioners in real time. This will require two persons: a play partner for the child and an observer to record the child's responses. Practitioners also will need effective training and the necessary administration materials for implementation, analysis, and interpretation to be easily accessible. As mentioned previously, additional work is needed to improve the sensitivity of the Movement GOMs for 2- and 3-year-olds. In addition, a larger sample of children that is monitored frequently over time will be necessary for establishing the representativeness and practical utility of these findings when used in the assessment of and intervention for young children. Finally, demonstrations that the Movement GOM is sensitive to specific forms of early intervention designed to accelerate progress toward the general outcome are needed to complete its development and to realize its potential. ◆

AUTHORS' NOTES

1. This research was supported by Grant H024560010, funded by the Early Education Programs for Children with Disabilities (EEPCD), Office of Special Education and Rehabilitation Services (OSERS), U.S. Department of Education. The opinions expressed herein do not necessarily reflect the position or policy of the U.S. Office of Education, and no official endorsement by the U.S. Office of Education should be inferred.
2. The authors would like to thank all the childcare centers, their staff, and the children and families who participated in the project. Thanks are due to Paige Bross, Judith Carta, Jodi Gager, Kari Manier, and Kenneth Parsley, who helped plan and carry out the work. In addition, thanks are due to the following persons, who provided invaluable assistance to this project: Deborah Linebarger, Jane Atwater, Debra Galvin, Dale Walker, and Eva Horn.

REFERENCES

Bayley, N. (1993). *Bayley scales of infant development* (2nd ed.). New York: Psychological Corp.

Bryk, A. S., & Raudenbusch, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods.* Newbury Park, CA: Sage.

Bryk, A. S., Raudenbusch, S. W., Cheong, Y. K., & Congdon, R. (2000). *HLM5: Hierarchical linear and nonlinear modeling.* Chicago: Scientific Software International.

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

Chandler, L. (1990). Neuromotor assessment. In E. M. Gibbs & D. Teti (Eds.), *Interdisciplinary assessment of infants: A guide for early intervention professionals* (pp. 45–61). Baltimore: Brookes.

Cratty, B. J. (1986). *Perceptual and motor developments in infants and children* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

Deno, S. L. (1997). Whether thou goest . . . Perspectives on progress monitoring. In J. W. Lloyd, E. J. Kameenui, & D. Chard (Eds.), *Issues in educating students with disabilities* (pp. 77–99). Mahwah, NJ: Erlbaum.

Deno, S. L., Mirkin, P. K., & Chaing, B. (1982). Identifying valid measures of reading. *Exceptional Children, 49*(1), 36–45.

Folio, M. R., & Fewell, R. R. (2000). *The Peabody developmental motor scales* (2nd ed.). San Antonio, TX: Psychological Corp.

Frankenburg, W. K., & Dodds, J. B. (1990). *Denver II screening manual.* Denver: Denver Developmental Materials.

Fuchs, L. S., & Deno, S. L. (1991). Paradigmatic distinctions between instructionally relevant measurement models. *Exceptional Children, 57,* 488–500.

Fuchs, L. S., & Fuchs, D. (1999). Monitoring student progress toward the development of reading competence: A review of three forms of classroom-based assessment. *School Psychology Review, 28,* 659–671.

Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1993). Formative evaluation of academic progress: How much growth can we expect? *School Psychology Review, 22*(1), 27–48.

Gallahue, D. L. (1989). *Understanding motor development: Infants, children, and adolescents* (2nd ed.). Carmel, IN: Benchmark Press.

Gallahue, D. L., & Ozmun, J. C. (1995). *Understanding motor development: Infants, children, adolescents, and adults* (3rd ed.). Madison, WI: Brown & Benchmark.

Gilfoyle, E. M., Grady, A. P., & Moore, J. C. (1981). *Children adapt.* Thorofare, NJ: Slack.

Goldfield, E. C. (1995). *Emergent forms: Origins and early development of human action and perception.* New York: Oxford University Press.

Greenwood, C. R., Luze, G. J., & Carta, J. J. (2002). Assessment of intervention results with infants and toddlers. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology* (Vol. 2, pp. 1219–1230). Washington, DC: National Association of School Psychologists.

Hannan, T. E. (1987). A cross-sequential assessment of the occurrences of pointing in 3- to 12-month-old human infants. *Infant Behavior and Development, 11,* 381–410.

Harris, S. R., & McEwen, I. R. (1996). Assessing motor skills. In M. McLean, D. B. Bailey, & M. Wolery (Eds.), *Assessing infants and preschoolers with special needs* (pp. 305–333). Englewood Cliffs, NJ: Merrill/Prentice Hall.

Hartmann, D. P. (1977). Considerations in the choice of interobserver reliability estimates. *Journal of Applied Behavior Analysis, 10,* 103–116.

Hatton, D. D., Bailey, D. B., Burchinal, M. R., & Ferrell, K. A. (1997). Developmental growth curves of preschool children with vision impairments. *Child Development, 68*(5), 788–806.

Kaminski, R. A., & Good, R. H. (1996). Toward a technology for assessing basic early literacy skills. *School Psychology Review, 25,* 215–227.

Kirshner, B., & Guyatt, G. H. (1985). A methodologic framework for assessing health indices. *Journal of Chronic Diseases, 38,* 27–36.

Kuntz, S. (2001). *Caregiver assessment of movement skills–Gross motor.* Kansas City, KS: University of Kansas, Juniper Gardens Children's Project.

Lewin, L. M., Hops, H., Davis, B., & Dishion, T. J. (1994). Multimethod comparison of similarity in school adjustment of siblings and unrelated children. *Developmental Psychology, 29*(6), 963–969.

Luze, G. J., Linebarger, D. L., Greenwood, C. R., Carta, J. J., Walker, D., Leitschuh, C., et al. (2001). Developing a general outcome measure of growth in expressive communication of infants and toddlers. *School Psychology Review, 30*(3), 383–406.

McConnell, S. R. (2000). Assessment in early intervention and early childhood special education: Building on the past to project into the future. *Topics in Early Childhood Special Education, 20,* 43–48.

McConnell, S. R., McEvoy, M. A., Carta, J. J., Greenwood, C. R., Kaminski, R., Good, R. H., et al. (1996). *Early childhood research institute on program performance measures: A growth and development approach.* Washington, DC: University of Minnesota, Minneapolis, MN: Proposal funded by the Early Education Programs for Children with Disabilities, Office of Special Education and Rehabilitation Services, U.S. Office of Education.

McConnell, S. R., McEvoy, M. A., Carta, J. J., Greenwood, C. R., Kaminski, R., Good, R. H., et al. (1998). *Research and development of exploring solutions assessments for children between birth and age eight.* (ERIC Document Reproduction Service No. ED452639)

Miller, L. J., & Roid, G. H. (1996). *The T.I.M.E.: Toddler and infant motor evaluation.* San Antonio, TX: Psychological Corp.

Mowder, B. A. (1997). Typical infant development. In A. H. Widerstrom, B. A. Mowder, & S. R. Sandall (Eds.), *Infant development and risk: An introduction* (pp. 25–60). Baltimore: Brookes.

Payne, V. G., & Isaacs, L. D. (1991). *Human motor development: A lifespan approach* (2nd ed.). Mountain View, CA: Mayfield.

Piper, M. C., & Darrah, J. (1993). *Motor assessment of the developing infant.* Philadelphia: Saunders.

Piper, M. C., & Darrah, J. (1994). *The Alberta infant motor scale (AIMS).* Philadelphia: Saunders.

Priest, J. S., McConnell, S. R., Walker, D., Carta, J. J., Kaminski, R. A., McEvoy, M. A., et al. (2001). General growth outcomes for children between birth and age eight: Where do we want young children to go today and tomorrow? *Journal of Early Intervention, 24*(3), 163–180.

Ramey, C. T., Breitmayer, B. J., & Goldman, B. D. (1984). Learning and cognition during infancy. In M. J. Hanson (Ed.), *Atypical infant development* (pp. 237–279). Baltimore: University Park Press.

Smith, K. U., & Smith, W. H. (1962). *Perception and motion.* Philadelphia: Saunders.

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action.* Cambridge, MA: MIT Press.