Infant Temperament: Actometer Measures and Developmental Correlates of Activity Level

by

Constance M. Dureski

A thesis
presented to the University of Manitoba
in fulfillment of the
thesis requirement for the degree of
Master of Arts
in
Psychology

December, 1991

National Library of Canada
Bibliothèque nationale du Canada

Acquisitions and Bibliographic Services Branch
Direction des acquisitions et des services bibliographiques

395 Wellington Street
Ottawa, Ontario
K1A 0N4

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

ISBN 0-315-77906-3

Canada

INFANT TEMPERAMENT: ACTOMETER MEASURES AND DEVELOPMENTAL CORRELATES OF ACTIVITY LEVEL

BY

CONSTANCE M. DURESKI

A thesis submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements of the degree of

MASTER OF ARTS

TABLE OF CONTENTS

LIST OF TABLES

ABSTRACT

Motor activity level (AL), a core dimension of child temperament, was cross-sectionally examined in a clinic setting using actometers and ratings on the Bayley Infant Behavior Record (IBR). Actometer measures of AL in infants aged 6 to 15 months were related to experimenter ratings on the IBR, objectively validating the IBR rating approach to assess AL. Examining the correlates of AL, actometer measures and IBR ratings of AL were positively related to IBR ratings of gross motor coordination, across all ages. While overall actometer scores were unrelated to mental and psychomotor development, several positive relationships emerged between rated AL and Bayley test performance. Rated AL was significantly correlated with Bayley Mental Development at 6 and 9 months, and with Bayley Psychomotor Development at 6, 9, and 12 months. Finally, cephalocaudal development was significantly correlated with actometer scores; arms were more active than legs. AL may well serve as a mediating agent that influences developmental progression.

ACKNOWLEDGEMENTS

I would like to acknowledge the support and assistance of several people who helped make this project possible. I thank my academic advisor, Dr. Warren Eaton, for his guidance and assistance at every stage of the project. His critical feedback was always delivered promptly and delicately, and it greatly improved the quality of each draft. I am also grateful to my other committee members, Drs. Sally Longstaffe and John Whiteley, for their helpful comments and suggestions. Special thanks also go Drs. Michael Moffatt and Sally Longstaffe for their support and enthusiasm in allowing me to conduct my thesis as part of their ongoing research. Additionally, I am grateful to the parents and infants who participated in the research and to Jeanne Besant and Helen Jones for their assistance with data collection.

INTRODUCTION

Individual behavioral differences that are not the result of experiential factors have been distinguished as characteristics of temperament (Hubert, Wachs, Peters-Martin, & Gandour, 1982). Such differences are evident in infants during the first weeks and months of life, and some may be expected to display continuity over the course of development, while others may be expected to change. The assessment of initial behavioral differences among infants must include identification and measurement of temperament, but an explicit and universally accepted definition of temperament has not been offered. Several researchers have offered different approaches to this construct.

Thomas and Chess (1977) suggested that temperamental individuality is well established by two or three months of age, and that the origins of temperament must be sought in genetic, prenatal, and early postnatal parental influences. They equated temperament with behavioral style. As the stylistic component of behavior, temperament refers to the how of behavior, as differentiated from the what (abilities and content) and the why (motivation) of behavior. The Thomas and Chess approach has been dominant, leading to the construction of numerous other parent-rating temperament questionnaires such as Carey's (1970) Infant Temperament Questionnaire (revised by Carey and McDevitt, 1978), Rothbart's (1981) Infant Behavior Questionnaire, the Behavioral Style Questionnaire (McDevitt & Carey, 1978), and the Middle Childhood Temperament Questionnaire (Hegvik, McDevitt, & Carey, 1982).

In a somewhat different approach, Goldsmith and Campos (1982a, 1982b) defined temperament in terms of emotional behavior. Their approach focuses on individual differences in the expression of what they consider to be primary emotions: motoric activity, anger, fearfulness, pleasure/joy and interest/persistence. Goldsmith and Campos' (1982a, 1982b) perspective on temperament is an example of the individual differences approach, derived from an interest in individual differences among infants and focused mainly on behavioral and emotional arousal. In measuring temperaments, Goldsmith and Campos (1982a, 1982b) focus on the intensive and temporal aspects of behavior. The emphasis on temporal aspects of behavior (e.g., time between stimulus and response, and time taken for response to reach maximum intensity) were borrowed from Rothbart and Derryberry (1981).

Rothbart and Derryberry (1981) moved to a psychobiological theory of temperament. They focused on reactivity and self-regulation as central properties of temperament. In defining temperament as relatively stable, biologically based, individual differences in reactivity and self-regulation, their approach attempts to link temperaments to the excitability or arousability of behavioral, endocrine, autonomic and central nervous system response (as measured by response threshold, latency, intensity, rise time, and recovery time). Self-regulation refers to processes, such as attention, approach, avoidance, and inhibition, which modulate (enhance or inhibit) reactivity (Goldsmith, Buss, Plomin, Rothbart, Thomas, Chess, Hinde, & McCall, 1987). These concepts have been operationalized through the measurement of temperament variables in infants: activity level, smiling and laughter, fear, distress to limitations, soothability, and duration of orienting (Rothbart, 1981). The Infant Behavior Questionnaire [IBQ] (Rothbart, 1981), a widely used parent report questionnaire, is one of a number of instruments that has been developed for the measurement of temperament, as part of the current interest in early individual differences.

Buss and Plomin (1975, 1984) took a different approach and argued that the most important criterion of temperament is inheritance, because this is what distinguishes temperament from other personality traits. The distinction between temperament and personality traits is important, despite the close resemblance between the two: Temperament is more applicable to the longitudinal study of stability in individual differences because temperament theorists have argued that stability and heredity are the two criteria which separate temperament from personality. Buss and Plomin (1984) emphasize that, in selecting those traits of temperament to investigate, it is important to consider their impact on personality development. Temperaments either show up continuously throughout development or leave behind residuals that determine the development of later personality traits. They define temperament as a set of inherited personality traits that appear early in life. Their theory suggests three temperaments: emotionality, activity, and sociability, offered as the acronym EAS. These three traits are considered to be broad, because they consist of behaviors that occur in most situations and must therefore be averaged across situations (activity) or behaviors that occur in repeated situations (emotionality and sociability). In the measurement of initial behavioral differences among infants, the present investigation adopted Buss and Plomin's (1975, 1984) approach to temperament as early-appearing inherited personality traits that resist modification.

Measurement of Temperament

Given current concern about the lack of agreement on a generally accepted theoretical definition of temperament, it is common practice to define temperament operationally, based on instruments or techniques used to assess this construct. In a review of temperament instruments, Hubert et al. (1982) suggest that great care be exercised in the choice of instruments used to assess infant or child temperament. Presently, the majority of approaches to assessing temperament rely either on parent interview, observation of infants and children, or parent questionnaires. Hubert et al. (1982) present a pattern of evidence indicating major reliability and construct validity problems in the measurement of temperament.

With the development of adequate measures of temperament, the questions of validity and reliability can be addressed. Assuming that temperament is a relatively stable characteristic of the organism (Bates, 1980), the lack of agreement between parent ratings, or between parents and observers, is hypothesized by Hubert et al. (1982) to be indicative of the weak operational status of current approaches to measuring infant and child temperament. What this means in practice is that there may well be ambiguities in exactly what is being rated or measured; parents and observers may identify different criteria for rating behaviors. Also, the use of a single assessment instrument or method, such as parent report, may result in findings which are specific to that instrument or methodology. The variability of reported results across studies is not surprising and, at the same time, the fact that empirical consistencies occur in spite of deficient measurement highlights the relevance of temperament as a developmental phenomenon.

A major goal must be to improve the measurement of temperament in infancy and childhood. Through improvement in measurement and assessment of individual differences in temperament, logic and data can more adequately test theories of temperament. In terms of future developmental research, greater emphasis should be placed on objectively defined measures, measures based on the observation of specific behaviors and/or on performance measures, within and across developmental periods. The present study applies such an objective measurement approach in the evaluation of a core dimension of infant temperament, activity level, across different infancy ages. By using an objective measure, the reliability and construct validity problems of rating instruments are bypassed.

Activity Level: A Core Component of Infant Temperament

One dimension of temperament that lends itself most readily to objective measurement is activity level (AL). Of all the theoretical dimensions of temperament, AL appears to be the dimension most commonly reported across studies (Matheny, 1980). AL, as Eaton and Enns (1986) define it, is the individual's customary level of energy expenditure through movement. In terms of causation, there is evidence for both genetic (Goldsmith & Gottesman, 1981; Matheny, 1980; Saudino & Eaton, 1991) and environmental (Eaton & Keats, 1982) components. The stability of AL has particularly interested researchers of infancy and early childhood.

Stability of AL

In longitudinal studies of temperament, stability refers to consistencies in behavior over time. Rothbart (1981) found significant stability in infant AL using her Infant Behavior Questionnaire across three-, six-, and nine-month periods. Studies by Halverson and Waldrop (1976) and Buss, Block, and Block (1980) have demonstrated stability in preschool AL over several years. Korner, Zeanah, Linden, Berkowitz, Kraemer, and Agras (1985) found in a longitudinal study using a parent rating scale and an electronic activity monitor that the vigor of neonatal movements was positively related to later preschool activity. Also, the least vigorous infants tended to become the least active preschoolers.

On the other hand, not all evidence for the stability of AL across time is confirmatory. Eaton and McKeen (1990) assessed AL in 6-month-old infants in the home environment over a 48-hour period using actometers and Rothbart's IBQ (1981). When they assessed AL again at 24 months of age using the Toddler Behavior

Assessment Questionnaire (Goldsmith, 1987), no relationship was found between six-month infant activity and subsequent toddler activity. However, because different measures were used at the two ages, the failure to find a relationship between infant and toddler activity may be the result of method variance rather than a true null relationship.

In a study of a similar age range, McDevitt and Carey (1981) found significant stability for all nine temperament dimensions outlined in the New York Longitudinal Study (Thomas, Chess, Birch, Hertzog, & Korn, 1963; Thomas & Chess, 1977) when they used caretaker ratings on the revised Infant Temperament Questionnaire (Carey and McDevitt, 1978) and the Toddler Temperament Scale (Fullard, McDevitt, and Carey, 1979). Unlike the Eaton and McKeen (1990) study, McDevitt and Carey's (1981) study relied solely on caretaker ratings. Thus, one cannot rule out the possibility that the stability they found was in the heads of the raters, not in the behavior of the infants.

The most extensive study of temperamental development from infancy through early childhood is the New York Longitudinal Study (Thomas, et al., 1963; Thomas & Chess, 1977). This research assessed the stability of nine temperamental variables (AL, rhythmicity, approach or withdrawal, adaptability, intensity of reaction, threshold of responsiveness, quality of mood, distractibility, attention span, and persistence) across the period from three months to two years of age. Greatest stability was found for mood, adaptability, approach behavior, and intensity (Thomas et al., 1963). A follow-up study, using the same temperamental variables during the first five years of life, obtained the greatest stability for activity, rhythmicity, and adaptability. Approach or withdrawal, distractibility, and persistence showed the least stability (Thomas & Chess, 1977).

Thomas et al. (1963) also reported that the first year's assessment was not predictive of behavior during the fifth year for any of the variables. As the time span between measurement of variables increased, the number of significant correlations decreased. This pattern of results implies that the age of the child at the time of measurement can affect correlational stability. Therefore, it is important to consider the developmental profile for temperamental dimensions. The increase or decrease of, say, AL, could clearly affect estimates of stability in individual diferences.

<u>Developmental Pattern</u>

In his review of 57 studies with information on age and activity in humans, Eaton (1991) found that a developmental curvilinear pattern emerged, characterized by an increase in AL during infancy, a peak between 2 and 5 years, and a decrease across the remainder of the lifespan. The developmental pattern of AL within the infancy period is examined in the present investigation. Motor AL was cross-sectionally examined in infants at 6, 9, 12, and 15 months of age, and following Eaton's (1991) findings, it was predicted that AL would increase across age, with AL at 15 months being greater than AL at 6 months. An emphasis was placed on the developmental pattern of AL and its relationship to cognitive performance within the infancy period.

Development and Temperament

In considering the developmental features of AL, it is difficult to ignore the simultaneous cognitive changes in the infant. The relationship between early-appearing personality traits and developmental test performance has been of interest to temperament researchers. In a study of the stability of newborn temperament and mother-infant interaction, Fish and Crockenberg (1981) report that five-day Brazelton motor-maturity [Neonatal Behavioral Assessment Scale] (Brazelton, 1973) correlated positively (.51, $p < .05$) with nine-month infant large motor activity (sitting unsupported, walking alone or holding onto person or object, rolling, creeping or crawling, pulling up, or standing). In terms of mother-infant interaction, the more motorically mature infants, as assessed by the Brazelton, had mothers who interacted less with them at nine months (.47, $p < .05$). Fish and Crockenberg (1981) interpret these findings to imply that "motorically well-developed infants may be more capable of organized motor activity, experience more success/pleasure in its expression, and thereby facilitate their own continued motor development" (p. 78).

Individual differences in motoric competency having AL antecedents in early infancy have also been reported by Matheny and Brown (1971). Infants reported by their mothers to be more active at age 12 months showed more advanced behaviors in locomotion and postural control, and walked without support significantly earlier than their less active co-twins. While rated as having inferior gross motor coordination and walking later, the less active twins had better fine motor coordination, were more attentive, and had a significantly higher Wechsler Primary and Preschool Scale of Intelligence (WPPSI) Performance IQ score when tested three years later, at age four. Thus, Matheny and Brown's (1971) findings indicate that the *less* active, more attentive infants persist longer at manipulative play, and they suggest that "this behavior, practiced from an early age, should facilitate the development of manipulative skills involving spatial relations, attention to figural-ground contrasts and discrimination among the properties of objects" (p. 156-157). Matheny and Brown (1971) report that the lack of agreement among studies that have explored the relationship between AL and differences in cognitive abilities is often attributed to "differences in the definition of activity, measurement methods, and the modifying effects of different familial situations" (p. 152).

Although Matheny and Brown (1971) report a significant relationship between low activity and better cognitive performance, interpretation of their results is limited because of the nature of the measures used in their study. Individual differences in temperament were measured in the absence of an established instrument, with the interviewer recording which twin was reported to display a behavior more prominently. Measures of AL and attention span in 12-month-olds were obtained using a parent interview, where mothers were asked to pinpoint similarities or differences between twins' behaviors. Matheny and Brown's (1971) interpretations are based on the assessment of a lagged relationship between mothers' reported perceptions of their twins' temperament at age 12 months, and the twins' WPPSI performance IQs, 3 years later.

Behaviors categorized as AL in Matheny and Brown's (1971) study may have been reflections of inattentiveness and/or impulsivity. While Matheny and Brown (1971) report that less active, more attentive children persist longer at manipulative play, Weithorn, Kagen, and Marcus (1984) have found that attention and impulsivity are more predictive of school success in the second grade than ratings of AL alone. Inaccurate maternal perceptions may have constituted the ratings of AL in Matheny and Brown's (1971) study. As Palisin (1986, p. 769) put it:

> These behaviors are not synonymous--the motorically active child is not necessarily impulsive, and the impulsive child may not be highly active. Nor does it follow that a child who is active will not be able to attend to intellectual tasks. One can think of children who will engage in gross motor play with abandon, then turn to a more sedentary pursuit such as reading with equal enthusiasm and concentration.

Though the longitudinal connection between AL and developmental test performance is ambiguous, the concurrent positive relationship between infant temperament and performance on the Bayley Scales of Infant Development (BSID;

Bayley, 1969) has been documented by Fagen, Singer, Ohr, and Fleckenstein (1987). Infants were assessed, at 4, 8, and 12 months of age, using the Rothbart Infant Behavior Questionnaire (Rothbart, 1981) and the BSID. AL was positively related to the infants' level of motor coordination displayed during the administration of the BSID at 8 and 12 months. In addition, maternal ratings of AL at 8 and 12 months were correlated with examiner ratings of infants' activity during the test administration, as measured by Bayley's IBR. Most prominently, of four temperament dimensions assessed in the Fagen et al. (1987) study (AL, fear, duration of orienting, and smiling and laughter), AL was the best and most consistent predictor of developmental status as measured by the BSID. Activity level was positively related to both mental and psychomotor development at every age tested. Fagen et al.'s (1987) findings suggest that infants may use environmental resources (objects in the world that the child can become engaged with through affect and cognition) in exploration and play, towards developmental advance.

Cognitive Correlates

Development in the 6- to 15-month age range is characterized by major cognitive changes, so the present study considered the possibility that developmental changes in AL may be linked to concurrent cognitive changes. For example, in addition to reflecting cognitive functioning per se, infant mental test scores may be related to dimensions such as AL. As age increases during infancy, increasingly sophisticated cognitive skills are reflected in increasing intelligence test raw scores. If AL increases during infancy, a positive correlation between AL and intelligence test raw scores is predicted. AL may well serve as a mediating, motivating agent that influences intellectual expression. Those children who are motivated to explore their environment are likely to be high in AL. The success and satisfaction inherent in exploration serves as intrinsic motivation to continue to explore and develop new skills, with AL influencing developmental progression. With this in mind, the present investigation assessed developmental test performance as a possible correlate of infant AL.

Actometer Measures of AL

In the study of individual differences in AL, an unusual opportunity for objective, unobtrusive measurement exists in the form of actometers. Actometers are motion recorders which, when attached to the limbs, are responsive to their movement. Eaton (1983) found these instruments to provide reliable and valid measures of preschooler AL; actometer readings correlated strongly with teacher and parent ratings of AL. Eaton and McKeen's (1990) results also provide evidence of a strong correlation between parent perception of AL and actometer readings of six-month-old infants.

On the other hand, Eaton and Dureski (1986) did not find a correlation between 24-hour actometer readings and parent ratings for a sample of three-month-olds. The purpose of the Eaton and Dureski (1986) investigation was to validate the activity scale of the Infant Behavior Questionnaire (IBQ) (Rothbart, 1981), which requires parents to rate the relative frequency of concrete behaviors in specific situations. With its emphasis on frequency counts of concrete behaviors in various situations, the IBQ activity scale should correlate positively with an aggregated actometer measure of the infant's limb movements. Rothbart (1986) similarly failed to find a significant correlation between IBQ Activity Level and observed limb movements for three-month olds, but found a significant correlation between observed limb movements and IBQ Activity at six and nine months.

As the preceding examples illustrate, the evidence on parent perceptions seems contradictory. Although he has argued that parent perceptions make an important contribution, Bates' (1980) suggestion that parent ratings of infant temperament represent parent perceptions rather than accurate reflections of behavior, may explain some of the incongruence between parent reports and other temperament measures and may underlie one of the disadvantages of using parent ratings in isolation. Rothbart and Derryberry (1981), on the other hand, argue for the utility of parent reports, while cautioning that they cannot represent a pure and independent measure of infant temperament.

Another plausible explanation for the contradictory data is that temperament changes with age (Wilson & Matheny, 1983). Wilson and Matheny (1983) explain Carey's (1981) position that there are shifts in temperament with age due to "varying rates of maturation for the underlying central nervous system structures, and the appearance of age-linked behavioral competencies" (p. 182). Perhaps such a maturational shift occurs between 3 and 6 months. If so, parent perceptions may be more accurate at a later age. Thus, the present study used four age groups, 6, 9, 12, and 15 months, to assess the validity of observer perceptions of infant AL.

A second reason for Eaton and Dureski 's (1986) failure to find a correlation between parent ratings on the IBQ and actometer readings may be due to an incomplete overlapping of parent and actometer samples of behavior. While the actometers recorded AL continuously over a 24-hour period, the parents answered questions about AL in specific contexts. It is necessary to assess the validity of rater perceptions within a context where the samples of behavior are the same for the observer and the actometer. One such context is the Bayley assessment situation, where the rater's or examiner's impressions are based on the same sample of behavior that the actometers record. The present investigation objectively measured individual differences in AL with actometers and, at the same time, examined Bayley Developmental status as a possible correlate of infant AL.

Bayley Scales of Infant Development

The Bayley Scales of Infant Development (1969) were standardized on a representative national sample of 1262 children in fourteen age groups from 2-months to 30-months-old. The Scales have been described as "the best measure of infant development available" with "excellent" norms and "satisfactory" reliability and validity (Sattler, 1982). The Scales were designed to evaluate a child's developmental status from the age of two months to two and one-half years. The Mental Scale was designed to assess sensory-perceptual abilities; early memory, learning and problem-solving abilities; vocalization and the beginning of verbal communications; and the beginnings of the ability to form generalizations and classifications for later abstract thinking. The Motor Scale was designed to provide a measure of body control, coordination of the large muscles and finer manipulatory skills of the hands and fingers.

A positive correlation between AL and Bayley Mental and Motor Development raw scores was expected. In addition, a positive relationship was expected between actometer measures and Bayley Infant Behavior Record (IBR) ratings of gross motor coordination and AL. Hence, the present investigation used actometers in an attempt to validate the IBR rating approach to assess AL.

Secondary Hypotheses

Sex Differences

The present study also considered sex differences as another correlate of AL during the infancy period. Eaton and Enns (1986) addressed the issue of sex differences in AL in their meta-analysis of 205 male-female comparisons from 127 studies. Their review supports the notion that sex differences exist, with males being more active, and the present investigation predicted the same, greater male activity.

Body Size

The role of temperament in infant body size has been considered, with particular attention to the role of the infant's activity-for-weight status. Rose and Mayer (1968) studied the relationship between AL, fat storage, and caloric intake in 31 infants between 4 and 6 months of age, and found an inverse relationship between subcutaneous fat and caloric intakes of the extremely fat and the extremely thin infants. Actometer-measured AL correlated -.53 with triceps skinfold thickness, a measure of body fat, and was also strongly related to total caloric intake (r=.47). The heavier infants expended fewer calories on activity, while thinner babies were more active.

In a similar investigation of the relationship between AL, growth, and caloric intake, Mack and Kleinhenz (1974) followed 5 infants from birth through the eighth week of life. The least active babies consumed more calories and gained weight faster than the most active babies. "The positive correlation between AL and gain in length (r=.91), together with the inverse correlation between AL and gain in weight (r=-.84) reflects the distribution of taller, more slender and more active infants at one extreme, and shorter, heavier, less active infants at the other" (p. 351). Eaton and Dureski (1986) also found that the thinner babies were more active in their investigation of infant motor AL. Actometer-measured activity of 3- to 4-month-old infants was significantly correlated with ponderal index (r=-.29). Ponderal index (PI), a weight-length ratio calculation of the amount of the infant's soft-tissue mass, was popularized by Miller and Hassanein (1971). Babies with more subcutaneous fat have a higher PI than babies with less subcutaneous fat.

To explore the relationship between neonatal and later activity and adiposity, Berkowitz, Agras, Korner, Kraemer, and Zeanah (1985) reassessed AL in a cohort of 52 children aged 4 to 8 years whose activity had been measured during the first 3 days of life. Electronically-monitored neonatal activity was not significantly correlated with adiposity at birth, nor did it predict adiposity in childhood. AL did, however, show longitudinal stability. In addition, children's daytime high activity (corresponding to activities at least as vigorous as walking or playing) was significantly associated with childhood adiposity. The absence of the activity-adiposity relationship at birth and a relationship at later ages in this and other studies (Rose & Mayer, 1968; Mack & Kleinhenz, 1974; Eaton & Dureski, 1986), prompted Berkowitz et al. (1985) to speculate that "physical activity becomes increasingly more important as a factor in determining adiposity as the child grows older. Because of the relationship between childhood activity and adiposity, we may speculate that there is a relationship between activity and the physiology of the fat cell during development" (p. 737). Failure to find any relationship between neonatal activity and neonatal or childhood adiposity may reside in Berkowitz et al's. (1985) measure of activity level. Neonatal physical activity was measured with an electronic activity monitor for 4 hours during the night and 4 hours during the day of one 24-hour period. The infants' activity during this time may not have been representative of activity level during later infancy.

In the present investigation, a negative correlation was expected between AL and ponderal index; thinner infants would be more active. Within a developmental perspective, babies who advance quickly would have a lower ponderal index (less subcutaneous fat) and would be both more active, and further along in Bayley Mental and Psychomotor status.

Also, in terms of developmental progression, it was expected that AL would be related to cephalocaudal development, with arms being more active than legs. This expected arm versus leg distinction in infant AL was found in Eaton and Dureski's (1986) investigation of motor AL in the 3- to 4-month-old.

Summary

Because rating instruments are commonly used to measure infant temperament, a joint strategy employing an objective, independent assessment of the child behavior upon which those rater perceptions are based was suggested. It was expected that this strategy would (a) make it possible to objectively measure individual differences in temperament (AL) as they relate to specific behavioral correlates (age, developmental competence, sex, body size), and (b) evaluate the adequacy or validity of the rating approach to assess temperament. In summary, the present research examined the relationships between mechanically-measured AL and several variables including age, Bayley Mental and Psychomotor developmental status, sex, and ponderal index. In addition, the validity of experimenter ratings of AL and gross motor coordination on the Bayley Infant Behavior Record (see Appendix A for activity and gross motor coordination items of the Bayley IBR) was addressed through objective verification. The following hypotheses were tested:

1) Actometer measures of AL will be positively related to IBR ratings of AL.

2) Age will be positively related to AL.

3) Mental Developmental Index will be positively related to AL.

4) Psychomotor Developmental Index will be positively related to AL.

5) Actometer measures of AL will be positively related to IBR ratings of gross motor coordination.

6) Boys will be more active than girls.

7) Ponderal index will be negatively related to AL.

8) AL will show cephalocaudal priority, with arms being more active than legs.

METHOD

Overview

Motor AL in infants and toddlers was assessed in a clinic setting using both a mechanical measure (actometer) and a rating measure (Bayley Infant Behavior Record). Moreover, several other variables such as chronological age were considered as possible correlates of AL. Among these variables were: developmental status, assessed using the Bayley Mental and Psychomotor Scales of Infant Development; various physical measurements taken during the infancy period, including length, weight, and head circumference; gestational age; birth weight; and Apgar scores. The young children were participants in an infant feeding study designed to investigate the effects of iron supplementation.

Sample Recruitment and Demographics

One hundred fifty one children participating in ongoing research in the Department of Pediatrics/Community Health Sciences, University of Manitoba comprised the base sample from which the study sample was drawn. Children in the base sample were those born to mothers who were registered in the "nonreferred" category at the Women's Centre, an obstetrics unit at the Health Sciences Centre. This category comprised women who were not under the care of a chosen or specified physician, and the babies born to these women were known as "staff babies." Many of the "staff babies" are monitored by the Children's Hospital Neonatal Follow-up

Clinic, which served as a second entry point for the study. A study coordinator recruited eligible participants at birth and from the Follow-up Clinic by identifying hospital charts that were tagged either "staff" or "nonreferred" and obtained written parental consent for enrollment in the study. Infants who had an obvious cause for potential developmental delay were excluded from enrollment.

The base sample was drawn from an inner city population predominantly Native in origin, and characterized by single, unemployed, young mothers living on social assistance. As part of the ongoing iron study, half of the participating infants received iron supplementation to their diet, and half did not. The effects of this intervention on the infants who also participated in the present study will be investigated at a later date, when the iron study is completed. At the present time, their iron status remains unknown because of the need for investigators to be blind to the individual infant's supplementation status.

Sample Derivation

Each child in the sample of 151 infants from the iron study was a potential candidate for four separate developmental/activity level assessments performed at 6, 9, 12, and 15 months during the infancy period. Only those assessments that included both satisfactory Bayley Scales and actometer assessments were included for consideration. An assessment was excluded if the Bayley assessment was judged as only a "minimal" or "fair" indicator of a child's characteristics (Bayley IBR item 28 scored as "1" or "2" by the examiner). Typically this occurred when the child presented as very tired, ill, fearful, irritable, or uncooperative at the time of testing.

Because the iron study from which this sample was drawn is longitudinal, many babies were assessed at more than one age. Only the first assessment that met the preceding criteria (e.g. complete actometer/developmental assessment) was included in the present cross-sectional sample. Assessments of children who had an obvious cause for potential developmental delay (i.e., cerebral palsy, fetal alcohol syndrome) were also excluded from the final sample and referred to the Child Development Clinic for further evaluation. Hence, including only those first-time assessments that involved both an actometer assessment of activity level and a Bayley developmental evaluation that was judged to be either an "average" or a "very good" indicator of the child's characteristics, the final sample included a total of 99 developmental/AL assessments involving children ranging in age between 5.7 and 15.8 months. Due to missed appointments and scheduling difficulties, the age of the children within each age group in the final cross-sectional sample varied somewhat. The distribution and basic summary statistics of the final cross-sectional sample according to age group is displayed in Table 1.

Physical Measurements

Following each AL/developmental assessment, measures of each child's weight and recumbent length were obtained from the research nursing staff, who routinely

Table 1. Summary Information of Cross-Sectional Sample, Overall and by Age Group

		Age Group (months)			
	Overall	6	9	12	15
Mean	9.7	6.3	9.6	12.3	15.3
SD	3.1	0.5	0.6	0.3	0.3
Range	5.7-15.8	5.7-7.6	8.7-11.4	11.8-12.9	14.8-15.8
n	99	34	33	20	12
Females	46	12	16	11	7
Males	53	22	17	9	5

assessed these growth parameters at each child's clinic visit. To measure the infant's length, the child was placed on his or her back on the firm, flat surface of the measuring board. A study coordinator assisted in straightening the infant's body. The infant's head was placed against the top of the board, and the sliding base of the board was aligned with the base of the infant's heel. The weight measurements were obtained by the same research staff, using the same digital scale with each child. Again, while the child was seated in the tray of the scale, a study coordinator assisted in steadying the infant's body. Measures of each child's length and weight provided the necessary information to assess the relationship between AL and ponderal index. The research nurse also obtained head circumference measurements using a measuring tape. A summary of the physcial measurements for the cross-sectional sample are presented in Table 2.

Perinatal Variables

Various measures that were taken at the birth of each participant in the cross-sectional sample were recorded from the infants' hospital charts. These measures included birth weight, gestational age, and Apgar scores at 1 and 5 minutes. An Apgar score is an expression of the neonatal status of a child, and is derived by evaluating five factors (heart rate, color, respiration, muscle tone, and reflexes) at 1 and 5 minutes after delivery. Each factor is rated on a 3-point scale (with 0 being weakest and 2 being strongest). The factor scores are summed to give a child an overall rating between 0 and 10. An extremely low score indicates that the newborn child may have a potential problem.

An antepartum high risk pregnancy score for each participant's mother was an additional measure obtained from the infants' hospital charts. The mother's score is derived by evaluating her pregnancy in relation to her reproductive history (including parity, previous long labour, neonatal death or difficult delivery), associated medical conditions (including previous gynaecological surgery, heart disease, or other medical disorders), and the present pregnancy (including bleeding, multiple pregnancy, breech, or malpresentation). Higher scores represent a greater risk that the mother and/or

Table 2. Physical Measurements of Infants in Cross-Sectional Sample

Age Group	M	SD	Range	n[a]
6 months				
Length (cm)	68.8	2.2	64 - 74	33
Weight (kg)	8.9	1.0	7 - 11	34
Head Circumference (cm)	44.6	1.6	42 - 47	24
Ponderal Index	2.7	0.2	2 - 3	33
9 months				
Length (cm)	72.7	2.2	69 - 77	31
Weight (kg)	10.1	1.2	8 - 12	32
Head Circumference (cm)	46.0	1.1	44 - 48	24
Ponderal Index	2.6	0.2	2 - 3	31
12 months				
Length (cm)	75.3	2.4	70 - 80	20
Weight (kg)	10.0	1.0	8 - 12	20
Head Circumference (cm)	46.3	1.8	43 - 49	11
Ponderal Index	2.3	0.2	2 - 3	20
15 months				
Length (cm)	80.0	2.3	77 - 84	12
Weight (cm)	11.7	1.4	10 - 14	12
Head Circumference (cm)	47.6	1.5	46 - 50	5
Ponderal Index	2.3	0.2	2 - 3	12

[a] unequal n's due to missing data.

newborn child may have a potential problem. The perinatal characteristics for the final cross-sectional sample are presented in Table 3. Because not all birth information was available to the research nurse, sample sizes vary in Table 3.

Developmental Evaluation

The developmental status of each child in the cross-sectional sample was assessed using the Bayley Mental and Psychomotor Scales of Infant Development. Abilities such as shape discrimination, purposeful manipulation of objects, sustained attention, imitation and comprehension, vocalization, memory, problem solving, and naming objects were assessed using the Mental scale. Gross and fine motor abilites such as rolling, sitting, standing, walking, and grasping were measured using the Motor Scale. Types of test items typical of each age group are illustrated in Table 4.

Assessment of AL

Rating measure. The Bayley Infant Behavior Record (IBR) rating scale was used to assess and record observations of each child's behaviors during the examination. To increase the reliability of activity-relevant IBR items, Matheny's (1983) factor analysis of the IBR was used to guide the aggregation of items. Four major clusters of items emerge during the 6 - 15 month age range: Task Orientation (i.e., goal directedness, object orientation, reactivity and attention span), Test Affect-Extroversion (i.e., cooperativeness, social responsiveness to examiner, emotional tone, endurance and fearfulness), Auditory-Visual Awareness (i.e., sensory interest in sights and sounds) and Activity (i.e., activity level, body motion and energy). Because the

Table 3. Perinatal Characteristics of Infants in Cross-Sectional Sample

Perinatal Variable	M	SD	Range	n[a]
Gestational Age (wks)	39.2	1.7	36 - 42	80
Birth Weight (gms)	3526.2	502.5	2590 - 4910	82
Apgar Score				
1 minute	7.7	1.7	1 - 9	81
5 minutes	8.8	0.5	7 - 9	81
Mother's Risk Score	2.4	1.9	0 - 8	52

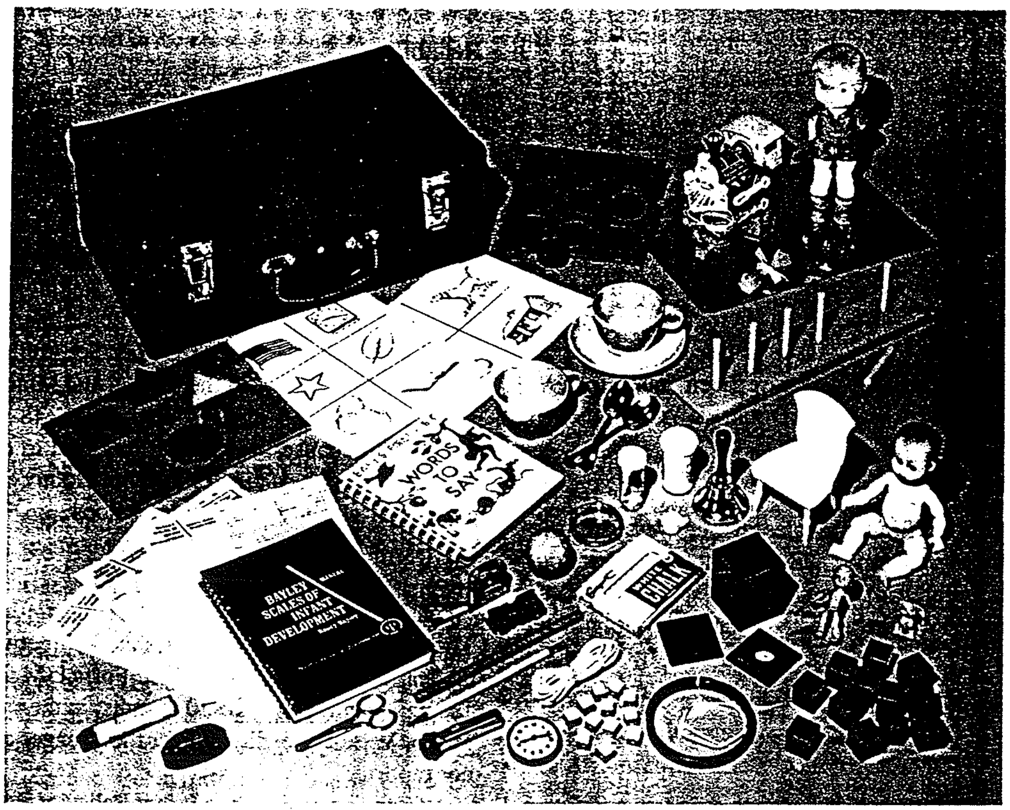
[a] unequal n's due to missing data.

hypothesis that the IBR rating of gross motor coordination would correlate positively with AL, the gross and fine motor items were examined separately, and not combined according to Matheny's (1980) motor coordination factor grouping. The IBR items for each factor were summed to create a unit-weighted composite for subsequent analyses, and the resulting IBR factors and their constituent items are displayed in Table 5. This composite score was used in all analyses involving hypotheses pertaining to IBR Activity.

Mechanical measure. Four Kaulins and Willis Model 101 motion recorders, or actometers, were used to obtain mechanical measures of activity. The actometer is a small conventional wristwatch that has been modified by unbalancing the pallet lever and removing the hair spring and balance wheel. Movement of the child's limbs and, hence, the actometers causes the second, minute, and hour hands of the actometers to advance, and the resulting differences between initial and final actometer readings in

Table 4. Illustrative Items from the Bayley Scales of Infant Development

Age (mos)	Mental Scale	Motor Scale
6	Looks for fallen spoon	Sits alone for 30 seconds or more
8	Uncovers toy	Pulls self to standing position
10	Looks at pictures in book	Walks with help
12	Turns pages of book	Walks alone
14	Spontaneously scribbles	Walks sideways
16	Builds tower of 3 cubes	Stands on left foot with help
18	Imitates crayon stroke	Tries to walk on walking board



[a] Adapted from Sattler, J. M. (1988).

total seconds are treated as activity units. The activity units registered during the

Table 5. IBR Factors and their Constituent Items

Factor & Items	Matheny Factor Loadings[a]			
	6 mo	9 mo	12 mo	18 mo
Activity				
Activity (#14)	.8	.8	.9	.9
Body Motion (#21)	.8	.8	.8	.9
Energy (#25)	.7	.5	.8	.8
Task Orientation				
Object Orientation (#8)	.8	.8	.8	.8
Goal Directedness (#11)	.8	.7	.7	.7
Attention Span (#12)	.7	.8	.8	.8
Reactivity (#15)	.7	.5	.5	.7
Test Affect-Extroversion				
Social- Examiner (#2)	.7	.6	.8	.7
Cooperativeness (#4)	.7	.7	.8	.7
Fearfulness (#5)	-.8	-.4	-.6	-.6
Emotional Tone (#7)	.7	.9	.9	.7
Endurance (#13)	.6	.8	.7	.6
Auditory-Visual Awareness				
Sights- Looking (#16)	.7	.6	.7	.8
Sounds- Listening (#17)	.7	.7	.6	.7

[a] From Matheny (1980), Tables 2-5.

recording interval were calculated using time formats included in SAS statistical software (SAS Institute, 1985). For example, an actometer with a reading of 10:36:07

upon attachment and a reading of 10:39:20 at removal produced a score of 193 activity units (AU's).

Watchbands were replaced with fabric bands with Velcro fasteners, which allowed for comfort and quick, convenient attachment and removal from the child's arms and legs. Actometers have been found to provide safe, reliable, and valid measures of both infant and preschooler activity level (Eaton, McKeen, and Lam, 1988; Eaton, 1983).

Summary of Developmental/Activity Assessment Procedure

The principal investigator of the present study administered the Bayley Scales of Infant Development to children ranging in age from 5.7 to 15.8 months. Each child's activity level was assessed in the testing setting, which was an examination room located in the Child Development Clinic at Children's Hospital. The room included an examining table, chairs for adults, a high-chair for the child, and, for children able to walk, free floor space with a path 10 feet long.

Upon arrival in the testing situation, the actual time of day was recorded, and four actometers were read, recorded, and placed on each of the child's four limbs in a predetermined order. Children assessed prior to 9 months of age were tested while seated in their caregiver's lap, and older infants were testing while sitting on their own, in a high-chair. Caregivers were asked to refrain as much as possible from restricting the limb movements of the infants who were seated in their lap. Infants were seated for the administration of the Bayley Mental Scale, and placed on the carpet for Bayley Motor testing. Following administration of the Bayley Scales, the actual time of day was again recorded (to calculate the length of the recording interval), and the four actometers were removed in the same order in which they were attached. The actometers were set aside where they remained undisturbed until the readings were taken. The mean length of the testing session was 30 minutes (SD = 10). The standardized assessment and recording procedure was the same for all children (see Appendix B for a sample of the data collection form).

The IBR was completed by the examiner immediately after administration of the Mental and Motor Scales. To eliminate the possibility of the experimenter being influenced by actometer scores, the instruments were read *after* the IBR had been completed. Inter-rater reliability was also assessed with an appropriately trained research assistant, who independently rated the behaviors of 25 subjects of varying ages.

No planned feedback was given immediately after data collection, but the experimenter answered any questions the participants had at this time. Information on each child's developmental progress was provided by the Iron Study staff to every parent when the final developmental assessment was completed.

RESULTS

Preliminary Analyses

Actometer Reliability

A composite measure of each child's activity was derived from the summed scores from the four actometers, and adjusted by minutes worn. The data for the interlimb and composite correlations is given in Table 6. The mean interlimb correlation, .40, provides an estimate of the reliability of a single actometer. The Spearman-Brown prophecy formula was then used to calculate the estimated reliabilty of the composite actometer score, which is based on four actometers. Thus calculated, the estimated reliabilty of the composite actometer score was .73.

Table 6. Actometer Intercorrelations by Limb and Composite for Cross-Sectional Sample

	Right		Left		
Limb	Arm	Leg	Arm	Leg	Composite
Right Arm		.19	.50 **	.18	.78 **
Right Leg			.16	.81 **	.58 **
Left Arm				.20 *	.78 **
Left Leg					.61 **

* $p < .05$. ** $p < .0001$.

As another reliabilty estimate, activity from the left and right sides of each child was correlated (i.e., scores from the actometers on the left arm and left leg were averaged and then correlated with the mean score from the actometers on the right arm and right leg). This right-left side correlation was .69, $p < .0001$.

Bayley Reliability

Interobserver agreement. Inter-rater reliability was assessed with an appropriately trained and experienced research assistant, who independently rated the behaviors of 25 subjects of varying ages at the beginning of the data collection period. Reliability coefficients for the BSID are presented in Table 7, where it can be seen that Pearson correlation coefficients between observers were .92 for both mental and psychomotor development.

Internal consistency. To increase the reliability of activity-relevant IBR items, Matheny's (1983) factor analysis of the IBR was used to guide the aggregation of activity items (i.e., activity level, body motion and energy). A unit-weighted composite measure of each child's rated activity was derived from the summed scores from the 3 activity items. The mean inter-item correlation, .88, provides an estimate of the reliability of a single IBR activity item. The Spearman-Brown prophecy formula was then used to calculate the estimated reliabilty of the composite activity score, which is based on 3 activity items. Thus calculated, the estimated reliabilty of the composite activity score was .96.

Table 7. Reliability Data for BSID

	r	Rater 1 mean (range)	Rater 2 mean (range)
MDI	.92	95.2 (54.0 - 147.0)	96.0 (54.0 - 150.0)
PDI	.92	97.2 (74.0 - 122.0)	95.3 (71.0 - 122.0)

$\underline{n} = 25$.

Convergent Validity

Of central interest was the correlation between the actometer score and the activity scale of the Bayley Infant Behavior Record. The results were strongly supportive of **Hypothesis 1:** Actometer measures of AL were positively related to IBR ratings of AL. Across all age groups: $\underline{r} = .56$, $\underline{p} < .005$. This relationship was significant for 6-, 9-, and 12-month age groups ($\underline{r}$'s = .50, .51, and .67 respectively) and of a similar magnitude at 15 months, $\underline{r} = .52$, n.s..

Age Changes

Analyses of variance with age group as the between-subject, independent variable, and actometer scores as the dependent variable indicated a significant between-subject effect, $\underline{F}$ (3,91) = 4.82, $\underline{p} < .004$, although not entirely in the predicted direction. Contrary to **Hypothesis 2**, that AL would increase with age, the present results indicate a "saw tooth" pattern between age and AL. This relationship is displayed in Table 8, where it can be seen that AL decreases between 6 and 9 months, increases between 9 and 12 months, and decreases again between 12 and 15 months.

AL Correlates

Overall AL as assessed by the actometers and by the experimenter was correlated with the following variables: Bayley Mental Development, Bayley Psychomotor Development, Bayley IBR gross motor coordination rating, sex, and

Table 8. Mean (SE) Activity Level, Overall and by Age Group

	Age Group (months)			
Overall	6	9	12	15
114.2 (4.6)	135.4 (9.8)	90.8 (6.3)	126.4 (9.4)	98.1 (9.4)

Ponderal Index. In order to test hypotheses 3 and 4, which predicted a positive relationship between AL and Bayley Mental and Psychomotor Development, the raw scores (uncorrected for age, representing the total number of successes achieved on the Mental and Psychomotor Scales) were used to analyze the AL relationship.

Mental Development

Hypothesis 3, that Bayley Mental Development would be positively related to AL as measured by the actometers, was unconfirmed for all age groups. However, Bayley IBR ratings of AL were significantly correlated with Bayley Mental Development at 6 and 9 months (r = .47 and .64 respectively, $p < .01$), as displayed in Table 9.

Psychomotor Development

Hypothesis 4, that Bayley Psychomotor Development would be positively related to AL, was supported only in part. While the overall relationship was not supported, several significant relationships emerged within individual age groups. Within the 9-month group, actometer scores were positively correlated with Bayley Psychomotor Development. Also, at 6, 9, and 12 months, IBR ratings of AL correlated positively with Bayley Psychomotor Development (see Table 9).

Table 9. Correlations of Activity Level and Bayley Development, by Age Group

Activity Level Correlate	Overall	6	9	12	15
	Age (months)				
Actometer Score					
Bayley Mental (raw score)	-.14	.08	.11	-.39	.28
Bayley Psychomotor (raw score)	-.10	.29	.35 *	.29	.19
IBR Rating					
Bayley Mental (raw score)	.00	.47 **	.64 **	-.43	.50
Bayley Psychomotor (raw score)	.03	.42 *	.55 **	.65 **	.18

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

Gross Motor Coordination

In order to test hypothesis 5, which predicted a positive relationship between AL and IBR gross motor coordination rating, the IBR gross motor item was reversed for the analysis, so that a high rating reflected smooth functioning and a low rating reflected poor coordination. Results were supportive of **Hypothesis 5**: AL was positively related to IBR ratings of gross motor coordination for all age groups combined for both actometer and IBR measures of AL (see Table 10). In an examination of the AL/gross motor coordination relationship according to age group, the relationship was sustained in the 9- and 12-month groups.

Sex Differences in AL

To test **Hypothesis 6**, that males would be more active than females, a repeated measures analyses of variance was conducted with sex as the between-subjects variable and the four limb actometer activity scores as the within-subjects variable. The results indicated that there were no differences in AL between males and females, $F\ (1,91) = 2.01$, n.s. In a separate correlational analysis, sex was also unrelated to experimenter ratings of AL on the Bayley IBR.

Ponderal Index

It was predicted, in **Hypothesis 7**, that ponderal index would be negatively related to AL. That is, babies with a lower ponderal index (less subcutaneous fat) were expected to be more active. Out of 8 possible correlations (ponderal index with IBR and actometer for each age group), the only evidence for this prediction was in the 12-month age group, where longer, thinner babies were more active as assessed by the actometers ($r = -.49$, $p < .03$).

Cephalocaudal Development

In terms of developmental progression, it was expected that AL would be related to cephalocaudal development, with arms being more active than legs. Using the four actometer limb scores as within subject variables and sex and age group (6, 9, 12, and 15 months) as between subject variables, a repeated measures analysis of variance was conducted. The results were strongly supportive of **Hypothesis 8:** AL

Table 10. Gross Motor Coordination Item Correlations with AL, by Age Group

Activity Measurement	Overall	Age (months) 6	9	12	15
Actometer Score	.32 *	.23	.46 *	.30	-.23
IBR Rating	.50 **	.28	.56 *	.75 *	-.01

* $p < .01$. ** $p < .0001$.

showed cephalocaudal priority, with arm readings significantly higher (M = 159.5) than leg readings (M = 68.8), $F(1,91) = 122.0$, $p < .0001$. Cephalocaudal development was also related to age group, $F(3,91) = 7.76$, $p < .0001$, with the arm versus leg difference being greatest at 6 months, and diminishing with age. Mean arm versus leg activity according to age group is displayed in Table 11. Cephalocaudal development also interacted with sex, $F(1,91) = 4.37$, $p < .04$, with the arm versus leg

Table 11. Mean (SE) Arm versus Leg Activity by Age Group

Age Group (mos)	Limb: Arm	Limb: Leg	Arm vs Leg Difference
6	200.7 (15.2)	70.1 (5.1)	130.6
9	136.3 (9.4)	45.3 (3.2)	91.0
12	164.6 (13.3)	88.2 (10.5)	76.4
15	98.4 (11.4)	97.8 (15.2)	0.6
Overall	159.5 (7.2)	68.8 (3.7)	90.7

difference greater in males (174.5 vs. 67.8) than in females (142.3 vs. 70.1).

Further examination of the limb scores in the same repeated measures analyses of variance revealed an interaction between cephalocaudal priority and laterality (left- vs. right-side activity), F (1,91) = 5.48, $p < .02$. This interaction shows that the arm-leg difference on the right side is greater than the arm-leg difference on the left. Mean limb activity for the cephalocaudal by laterality interaction is displayed in Table 12.

Table 12. Mean (SE) Limb Activity of Cephalocaudal by Laterality Interaction

	Left	Right	Left-Right Difference
Arm	156.1 (10.0)	163.0 (10.3)	-6.9
Leg	77.7 (5.9)	60.0 (4.3)	17.7
Arm-Leg Difference	78.4	103.0	

Laterality was also related to age, F (3,91) = 3.74, p <.01, with AL on the left side of the body greater than right-side AL at 6 and 9 months, and right-side AL greater than left-side AL at 12 and 15 months. Mean left- vs. right-side activity according to age group is displayed in Table 13.

Table 13. Mean (SE) Left- versus Right-Side Activity by Age Group

Age Group (mos)	Limb		Left vs Right Difference
	Left	Right	
6	149.3 (13.8)	121.6 (13.8)	27.7
9	91.8 (8.3)	89.7 (9.7)	2.1
12	116.4 (12.2)	136.4 (14.4)	-20.0
15	94.9 (14.0)	101.3 (12.8)	-6.4
Overall	116.9 (6.4)	111.5 (6.7)	5.4

DISCUSSION

Motor AL as measured by the actometers in a clinic setting was related to experimenter ratings on the Bayley IBR. Hence, through objective verification, the present research clarified the adequacy and established the validity of the IBR rating approach to assess AL, a core component of infant temperament. Furthermore, the developmental progression of AL was clarified through the evident relationship between AL and cephalocaudal development, with arms being more active than legs. Within a developmental perspective, it was found that AL decreased between 6 and 9 months, increased between 9 and 12 months, and decreased again between 12 and 15 months. While overall gross motor coordination was positively related to AL, the overall relationship between AL and Bayley Mental and Psychomotor status was not supported. Each of the results for the AL measurement and developmental correlate issues are discussed below.

Measurement of AL

The purpose of comparing mechanical and rating measures of AL was to address the operational status of the rating approach. Through improvement in measurement, the questions of reliability and validity could be addressed. The present research found that a single actometer has moderate reliability, while four combined actometers have reasonably good reliability for assessing variation in motor AL in young children, even though the assessment period was of a half hour's duration.

When the overall actometer score for each child was correlated with the Bayley Infant Behavior Record activity score, it was found that AL as measured by the actometers was related to experimenter ratings of activity on the Bayley Infant Behavior Record. This relationship was significant across all age groups, and also within the individual 6-, 9-, and 12- month age groups. Although of a similar magnitude at 15 months, the relationship was nonsignificant. The small sample of 12 children in the 15-month age group was a shortcoming of the present research. In view of the small sample size, the findings for this age group are unstable, and the remainder of the discussion will focus on the 6-, 9-, and 12-month age groups.

<u>AL Correlates</u>

The predicitons regarding overall AL and mental and psychomotor development were unsupported. That is, overall AL as assessed by the actometers and by the experimenter was unrelated to Bayley Mental and Psychomotor Scale raw scores. In addition, contrary to Eaton's (1991) finding that AL increases during the infancy period, the present study found that AL *decreased* between 6 and 9 months, and increased between 9 and 12 months. In the present investigation, the absence of a direct positive relationship between age and AL precluded a positive correlation between overall AL and developmental test raw scores. As age increases during infancy, increasingly sophisticated cognitive skills are reflected in increasing developmental test raw scores. Hence, with an expected increase in AL during infancy, a positive correlation was also expected between overall AL and developmental test raw scores.

Nonsignificant results also emerged for the hypothesized associations between AL and body size, and between AL and sex. Aside from a single negative correlation within the 12-month group, where longer, thinner babies were more active, overall there were no differences in AL between longer, thinner infants and shorter, fatter infants. In addition, contrary to Eaton and Enns' (1983) finding that boys under one year of age are more active than girls, the present study found no differences in AL between males and females. Also contrary to Eaton and Enns' (1983) finding of increasingly large sex differences with age (as environmental influences magnify existing differences), a sex by age interaction did not emerge in the present research. Several explanations for this lack of predicted AL correlations are possible.

Measurement Problems

Atypical experimental setting. In view of the constrained measurement situation, this lack of expected findings may be due to an unrepresentative sample of AL. Samples of AL were obtained in a somewhat specialized half-hour experimental session, where infants were exposed to a stranger posing a series of tasks. With its novel stimuli and short duration, the task setting may have been too atypical and too brief for the expected correlates of AL to emerge. Activity measured during the half-hour experimental session may have been unrepresentative of the infant's typical AL.

Stranger anxiety. Of all the infants in the present study, only those in the 6-month age group appeared to be unaware that actometers had been placed over the clothing on their limbs. After 6 months of age, this presumed lack of awareness seemed to disappear. With infants older than 6 months, differences in the rate of adaptation to actometer stimulation may account for the lack of differences in AL between males and females and for the decrease in AL between 6 and 9 months. Fear of strangers develops at about the same age in infants, regardless of cultural back ground (Dworetzky, 1981). This fear begins to occur after 6 months of age and peaks at about 9 months (Schaffer and Emerson, 1964). In the present study, stranger anxiety after 6 months of age may account for the decrease in AL between 6 and 9 months and subsequent increase in AL between 9 and 12 months. AL should ideally be assessed with longer behavioral samples and include a fixed adaptation period.

Iron status. An alternative explanation for the lack of a direct positive relationship between age and AL concerns the unknown iron status of the study participants. Because the infants in the present investigation were also participating in an iron supplementation study, their behavior may not be characteristic of unselected samples of infants. In their Guatemalan study, Lozoff, Klein, and Prabucki (1986) analyzed the behavior of iron-deficient anemic and nonanemic 6- to 24-month-old infants and their mothers from a videotaped free play session. Anemic infants initiated and maintained more body contact with their mothers than did nonanemic babies: 71 percent of the anemic group were high in body contact, compared to 29 percent of the nonanemic babies. Mothers of the anemic infants spent less time at a distance from them, were less likely to break close contact, and were more likely to re-establish close contact if the baby moved away. The body contact observed was interpreted to be a subtle manifestation of disturbance in affect, energy, or activity. An unknown half of the infants in the present investigation may have behaved atypically because of low levels of iron. In addition, the sample composition by age group in the present study may not be random with respect to iron supplementation, leading to an overall pattern of activity inconsistent with Eaton's (1991) documented pattern of increasing AL from the early months of fetal life to sometime in the preschool years.

AL and Development

The finding that actometer variance was significantly related to cephalocaudal priority of arms before legs parallels the results found by Eaton and Dureski (1986). Though part of the effect in the present investigation could be due to task demands being greater on arms than on legs, the evident relationship between AL and cephalocaudal development clarifies the developmental progression of AL. Cephalocaudal development was also related to age group, with the arm versus leg difference being greatest at 6 months, and diminishing with age. Increasingly sophisticated motor development in the lower part of the body is reflected in the onset of crawling and walking, and a diminution of arm versus leg AL differences with age.

In terms of motor and mental development, several significant relationships emerged within individual age groups between AL and Bayley Psychomotor and Mental Scale raw scores. Within the 9-month group, actometer scores were positively correlated with Bayley Psychomotor Development. IBR ratings of AL were significantly correlated with Bayley Mental Development at 6 and 9 months, and with Bayley Psychomotor Development at 6, 9, and 12 months. There were generally higher correlations of IBR activity ratings with Bayley performance as compared to actometer activity scores with Bayley performance, and this could be partly due to having the same person doing both (rating activity level and scoring the Bayley assessments). Nevertheless, the positive correlations between rated AL and Bayley test performance within individual age groups suggest that in addition to reflecting mental and psychomotor functioning per se, infant developmental test scores are related to AL. Furthermore, the correlations between actometer scores and Bayley Psychomotor performance are similar in magnitude to those obtained by Fagen et al. (1987) between rated AL on the Infant Behavior Questionnaire (Rothbart, 1981) and Bayley Psychomotor Index.

Additionally, AL was positively related to IBR ratings of gross motor coordination across all age groups. This relationship was significant for actometer measures and IBR ratings of AL. In an examination of the AL/gross motor coordination relationship according to age group, the correlations are once again similar to those obtained by Fagen et al. (1987) between rated AL on the IBQ

(Rothbart, 1981) and Bayley IBR ratings of gross motor coordination. While further studies are needed to establish the generalization of the IBR validity outside of the Bayley assessment situation, these consistencies across studies suggest that the present results are not entirely specific to the methodology. Gross motor coordination and general motor activity seem to be related in infancy.

Conclusions

In summary, the evident relationships between overall AL and gross motor coordination and between rated AL and psychomotor development at 6, 9, and 12 months of age suggest that AL may well serve as a mediating agent that influences developmental progression. Active children explore their environment, experience success and satisfaction inherent in exploration, and continue to explore and develop, with AL influencing gross motor coordination and psychomotor development. Even in view of the constrained measurement situation, the occurrence of empirical consistencies across studies suggests that the present findings are not entirely specific to the methodology or the instruments.

Furthermore, the actometer is a reliable and valid instrument for measuring variation in infant AL. The most important observation to be made from this study is that motor AL as measured by the actometers in a clinic setting was related to experimenter ratings of activity on the Bayley Infant Behavior Record. Hence, the present research objectively validated the IBR rating approach to assess AL and the issue of accuracy in rater perception of a core component of infant temperament has been clarified.

Directions for future research

Future examinations of the behavioral correlates of AL should attempt to assess activity with longer samples of behavior in less constrained situations, to obtain more representative samples of behavior. Inclusion of a fixed adaptation period may be necessary with older children. Ideally and if practical, two or more types of temperament measures should be used within the same study. Empirical consistencies in research that includes well-validated, reliable parent-rating instruments in addition to mechanical measures and experimenter ratings would further highlight the relevance of temperament to infant development more generally.

REFERENCES

Bates, J. E. (1980). The concept of difficult temperament. Merrill-Palmer Quarterly, 26, 299-319.

Bayley, N. (1969). The Bayley scales of infant development. New York: Psychological Corporation.

Berkowitz, R. I., Agras, W. S., Korner, A. F., Kraemer, H. C., & Zeanah, C. H. (1985). Physical activity and adiposity: A longitudinal study from birth to childhood. Journal of Pediatrics, 106 (5), 734-738.

Brazelton, T. B. (1973). Neonatal behavioral assessment scale. London: Spastics International Medical Publications.

Buss, D. M., Block, J. H., & Block, J. (1980). Preschool activity level: Personality correlates and developmental implications. Child Development, 51, 401-408.

Buss, A. H., & Plomin, R. (1975). A temperament theory of personality development. New York: Wiley.

Buss, A. H., & Plomin, R. (1984). Temperament: Early appearing personality traits. Hillsdale, NJ: Erlbaum.

Carey, W. B. (1970). A simplified method for measuring infant temperament. Journal of Pediatrics, 77, 188-194.

Carey, W. B. (1972). Clinical applications of infant temperament measurements. Journal of Pediatrics, 81, 823-828.

Carey, W. B. (1981). The importance of temperament-environment interaction for child health and development. In M. Lewis & L. A. Rosenblum (Eds.), The uncommon child. New York: Plenum Press.

Carey, W. B., & McDevitt, S. C. (1978). Revision of the infant temperament questionnaire. Pediatrics, 61, 735-739.

Dworetzky, J. P. (1981). Introduction to child development. Minnesota: West Publishing Company.

Eaton, W. O. (1983). Measuring activity level with actometers: Reliability, validity, and arm length. Child Development, 54, 720-726.

Eaton, W. O. (1991). Temperament, development, and the big five: The case of activity level. Paper presented at the conference on The Development of the Structure of Temperament and Personality from Infancy to Adulthood, Wassenaar, Netherlands.

Eaton, W. O., & Dureski, C. M. (1986). Parent and actometer measures of motor activity level in the young infant. Infant Behavior and Development, 9, 383-393.

Eaton, W. O., & Enns, L. R. (1986). Sex differences in human motor activity level. Psychological Bulletin, 100, 19-28.

Eaton, W. O., & Keats, J. G. (1982). Peer presence, stress, and sex differences in the motor activity levels of preschoolers. Developmental Psychology, 18, 534-540.

Eaton, W. O., & McKeen, N. A. (1990). Infant temperament: Longitudinal instability. Paper presented at the International Conference on Infant Studies, Montreal, Canada.

Eaton, W. O., McKeen, N. A., and Lam, C-S. (1988). Instrumented motor activity measurement of the young infant in the home: Validity and reliability. Infant Behavior and Development, 11, 375-378.

Fagen, J. W., Singer, J. M., Ohr, P. S., and Fleckenstein, L. K. (1987). Infant temperament and performance on the Bayley Scales of Infant Development at 4, 8, and 12 months of age. Infant Behavior and Development, 10, 505-512.

Fish, M. & Crockenberg, S. (1981). Correlates and antecedents of nine-month infant behavior and mother-infant interaction. Infant Behavior and Development, 4, 69-81.

Fullard, W., McDevitt, S., and Carey, W. (1979). The toddler temperament scale. Unpublished test form, Temple University, Philadelphia.

Goldsmith, H. H., & Campos, J. J. (1982a). Toward a theory of infant temperament. In R. N. Emde & R. Harmon (Eds.), The attachment affiliative systems. New York: Plenum Press.

Goldsmith, H. H., & Campos, J. J. (1982b). Genetic influence on individual differences in emotionality. Infant Behavior and Development, 5, 99.

Goldsmith, H. H., & Gottesman, I. I. (1981). Origins of variation in behavioral style: A longitudinal study of temperament in young twins. Child Development, 52, 91-103.

Goldsmith, H. H., Buss, A. H., Plomin, R., Rothbart, M. K., Thomas, A. Chess, S., Hinde, R. A., & McCall, R. B. (1987). Roundtable: What is temperament? Four approaches. Child Development, 58, 505-529.

Halverson, C. F., & Waldrop, M. F. (1976). Relations between preschool activity and aspects of intellectual and social behavior at age 7 1/2. Developmental Psychology, 12, 107-112.

Hegvik, R. L., McDevitt, S. C., & Carey, W. B. (1982). The middle childhood temperament questionnaire. Developmental and Behavioral Pediatrics, 3, 197-200.

Hubert, N. C., Wachs, T. D., Peters-Martin, P., & Gandour, M. J. (1982). The study of early temperament: Measurement and conceptual issues. Child Development, 53, 571-600.

Korner, A. F., Zeanah, C. H., Linden, J., Berkowitz, R. I., Kraemer, H. C., & Agras, W. S. (1985). The relation between neonatal and later activity and temperament. Child Development, 56, 38-42.

Lozoff, B., Klein, N. K., & Prabucki, K. (1986). Iron-deficient anemic infants at play. Developmental and Behavioral Pediatrics, 7, 152-158.

Matheny, A. P., Jr. (1980). Bayley's Infant Behavior Record: Behavioral components and twin analysis. Child Development, 52, 1157-1167.

Matheny, A. P., Jr. (1983). A longitudinal twin study of stability of components from Bayley's Infant Behavior Record. Child Development, 54, 356-360.

Matheny, A. P., & Brown, A. M. (1971). Activity, motor coordination and attention: Individual differences in twins. Perceptual and Motor Skills, 32, 151-158.

McCall, R. B. (1977). Challenges to a science of developmental psychology. Child Development, 48, 333-344.

McDevitt, S. C., & Carey, W. B. (1978). The measurement of temperament in 3 to 7 year old children. Journal of Child Psychology and Psychiatry, 19, 245-253.

McDevitt, S. C., & Carey, W. B. (1981). Stability of ratings vs. perceptions of temperament from early infancy to 1-3 years. American Journal of Orthopsychiatry, 51, 342-345.

Miller, H. C., & Hassanein, K. (1971). Diagnosis of impaired fetal growth in newborn infants. Pediatrics, 48, 511-522.

Palisin, H. (1986). Preschool temperament and performance on achievement tests. Developmental Psychology, 22, 766-770.

Rothbart, M. K. (1981). Measurement of temperament in infancy. Child Development, 52, 569-578.

Rothbart, M. K. (1986). Longitudinal observation of infant temperament. Developmental Psychology, 22, 356-365.

Rothbart, M. K., & Derryberry, D. (1981). Development of individual differences in temperament. In M. E. Lamb & A. L. Brown (eds.), Advances in developmental psychology (Vol. 1, p. 37-86). Hillsdale, N.J.: Erlbaum.

SAS Institute (1985). SAS user's guide: Basics (Version 5 ed.). Cary, NC: Author.

Sattler, J. M. (1988). Assessment of children. (3rd ed.) San Diego: Jerome M. Sattler, Publisher.

Saudino, K. J., & Eaton, W. O. (1991). Infant temperament and genetics: An objective twin study of motor activity level. Child Development, in press.

Schaffer, H. R., & Emerson, P. E. (1964). The development of social attachments in infancy. In K. Sherrod, P. Vietze, & S. Friedman, (1978), Infancy. California: Wadsworth Publishing Company.

Thomas, A., & Chess, S. (1977). Temperament and development. New York: Bruner/Mazel.

Thomas, A., Chess, S., Birch, H. G., Hertzog, M. E., & Korn, S. (1963). Behavioral individuality in early childhood. New York: New York University Press.

Weithorn, C. A., Kagan, E., & Marcus, M. (1984). The relationship of activity level ratings and cognitive impulsivity to task performance and academic achievement. Journal of Child Psychology and Psychiatry, 25, 587-606.

Wilson, R. S., & Matheny, A. P. (1983). Assessment of temperament in infant twins. Developmental Psychology, 19, 172-183.

APPENDIX A

Bayley Activity and Gross Motor Coordination Items

INSTRUCTIONS: Fill out this form immediately after the Mental and Motor Scales have been administered. For each rating scale, circle the number next to the one statement that best describes the child's behavior.

ACTIVITY RATING

14. Amount of gross bodily movement (Circle one)

1 Stays quietly in one place, with practically no self-initiated movement
2 Between 1 and 3
3 Usually quiet and inactive but responds appropriately in situations calling for some activity
4 Between 3 and 5
5 Moderate activity; enters into games with freedom of action
6 Between 5 and 7
7 In action during much of the period of observation
8 Between 7 and 9
9 Hyperactive; cannot be quieted for sedentary tests

	None						Excessive		
21. Body motion	1	2	3	4	5	6	7	8	9

25. Level of energy

Low				High
1	2	3	4	5

COORDINATION FOR AGE (Circle one)

Note.--Ratings for these items should be estimated in relation to other children of the child's own age.

26. Coordination of gross muscle movements

Smooth				Poor
1	2	3	4	5

APPENDIX B

Actometer Record Form

Name:__________________ ID:____

Date Tested (DDMMYY)____ ____ ____

Date of Birth (DDMMYY)____ ____ ____

Testing Session

Begin (hh:mm) ___ ___: ___ ___ End (hh:mm) ___ ___: ___ ___

Set ___ Acto #	Acto Start (hh : mm : ss)	Acto Stop (hh : mm : ss)
Right Arm	__ __: __ __: __ __	__ __: __ __: __ __
Left Arm	__ __: __ __: __ __	__ __: __ __: __ __
Right Leg	__ __: __ __: __ __	__ __: __ __: __ __
Left Leg	__ __: __ __: __ __	__ __: __ __: __ __

Interviewer: ______________________

Length: __ __.__ cm Weight: __ __ __ __ gm

Notes: