Aims and Objectives

The present study was conducted to construct a standardized Development Battery for Indian children aged 0 to 7 years and 11 months in five key areas of development. The specific objectives of the study were:

1. To construct a multi-domain development assessment battery for Indian children aged 0 to 7 years and 11 months.

2. To establish norms for children aged 0 to 7 years and 11 months in terms of domain and total test developmental quotient, and percentile scores.

3. To establish inter-rater, test-retest, and internal consistency reliability of the Development Battery.

4. To establish content, construct, and criterion related validity of the Development Battery.

Construction of the Development Battery

The development of the battery was done in three major stages: planning, pilot, and standardization phase. In the planning phase an extensive review of literature was done and five major areas of child development were identified: motor, adaptive, social-emotional, cognitive, and communication. In addition, each domain was further divided into two sub-domains. The sub-domains included, fine and gross motor for the Motor domain; self-care and personal responsibility for Adaptive domain; interpersonal interaction and social role for the Social-Emotional domain; attention and perception, and concepts and academics for the Cognitive domain; and receptive and expressive communication for the Communication domain. An extensive item bank was assembled, consisting of important milestones from the existing screening tests and batteries, books on child development, and a review of literature for targeted areas not addressed by existing tests. Further, the items were assigned to an approximate age level. A kit of manipulatives was assembled for administration of the
various tasks in the Development Battery. For each item, a standard procedure for assessing the skill or behavior was defined.

After assembling an item bank, consisting of 425 items, experts from the areas of child development, developmental pediatrics, and child psychology reviewed all the items. Each item was evaluated on its importance as a developmental milestone, ease of administration, and clarity of instructions. Based on the feedback from the experts, the items were modified.

In the pilot phase all the items from the assembled item bank were administered to 236 children recruited from immunization clinics, play way schools, and public and government schools of Chandigarh. On field testing, some of the items were found to be problematic in terms of their difficulty level, age appropriateness, ease of administration, negative reactions from children tested, or problems with the test materials. These items were reviewed and an effort was made to either adjust the item, and if that was not possible, the item was replaced with another item. In addition, the ambiguously worded items were rewritten or replaced. In cases where the difficulty level of the item was misaligned, the item was repositioned. Of the 425 items that were pilot tested, 13 items were deleted after the pilot study, and a total of 412 items were retained for the next phase of the study.

In the initial part of the Standardization study, the battery was administered to 626 children. The sample was stratified on the basis of age, sex, and socio-economic status. The sample comprised of 18 age groups. Children in the first 2 years of age were divided into 4 month groupings; children 2 to 7 years and 11 months were divided into 6 month groupings. A total of 366 boys and 260 girls were recruited. The battery was designed in a manner that some items could be scored by observation, direct elicitation of the response, or interview of the parent or caregiver.

An objective three-point (2, 1, and 0) scoring system was devised for the test battery. This scoring system was adopted to take into account emerging as well as fully developed skills. A score of “2” was assigned if the child consistently performed the task or if the milestone had been achieved. A score of “1” was assigned when the skill being tested was still emerging and the child could perform the task only about
20 to 50% of the time. Any skill which the child performed inconsistently (<20% of the time) was deemed as a future milestone and that item was scored as "0".

Rules for establishing the starting point, basal and ceiling levels were clearly defined to make the testing process more efficient. With a child of normal ability, the starting point was set at two levels below the chronological age. For all sub-domains, the basal level was reached when the child scored "2" on four consecutive items. If the child did not score "2" on four consecutive items, the items were administered in reverse order until a score of "2" was received on four consecutive items. The ceiling level was reached when the child scored "0" points on four consecutive items. Once the basal had been established, testing was continued until the child scored "0" on four consecutive items.

All data from the standardization phase test administrations were checked for accuracy and scored. Extensive analyses were conducted on the data including conventional item analysis, studies of item difficulty across age groups, item gradients, factor analyses, studies of differential item functioning and test fairness across gender and socio-economic status groups.

Each item was placed at the age where 75% of children passed that item within the given age group. The 75% criterion was selected because it represented a clear majority of children in a given age range while recognizing the variation in individual developmental rates. In order to meet the goals of non-discrimination as a function of sex and residence and socioeconomic status, the percentages of girls and boys; and children from rural and urban backgrounds and from lower, middle and upper socioeconomic backgrounds were compared. Items with highly discrepant passing rates (e.g., 40%) were eliminated. Items with slight discrepancies were "balanced" within age levels. Issues pertaining to item reliability within the context of the sub-domains were studied. This included the differentiation of responses to the item from high and low ability children, and the clarity of the scoring criteria. Based on this, new scoring criteria were written for some items.

The final selection of the items for the standardization phase comprised various sources of information, item analyses and the relative merit of items and sub-domains from the pilot phase of the study. Some of the criteria employed in the final selection
of the items were as follows:

1. High rating by experts involving criteria such as importance of milestone, ease of administration and the child's responsiveness.

2. Relative freedom from gender, socio-economic or cultural bias.

3. Item readily amenable to objective scoring procedures.

4. Each item contributed uniquely and positively to the domain structure and the total test.

The application of the above mentioned criteria resulted in the selection of the items which appear in the final version of the test battery. A total of 10 sub-domains provide measure of the five domains of the test: Adaptive, Social-emotional, Communication, Cognitive and Motor. Out of the 412 items initially included in the standardization phase, only 378 items were retained for the final version of the test and the scores from these 378 items were used in the normative data analysis (phase 2 of the standardization study).

Development of the Norms

Domain standard scores: A procedure devised by Angoff and Robertson (1983) was used to develop normalized domain standard scores ($M=100, SD=15$) for the test battery. The same procedure has been used to determine standard scores for the Kaufman Assessment Battery for Children (K-ABC; Kaufman and Kaufman, 1983) and the Vineland Adaptive Behaviour Scales (Sparrow et al., 1984). The steps in the development of norms were:

1. The standardization sample of 626 children was divided into 18 age groups. For each age domain, the the raw score means and standard deviations for all the 18 age groups were plotted on arithmetic graph paper, and the smoothed means and standard deviations were read from the curves drawn through the plotted points. The smoothed means and standard deviations were used in subsequent steps.

2. For each domain, one age group whose mean raw score was near the center of the total raw score distribution was selected as the anchor age group. Raw
score distributions for the remaining age groups were converted to the designated anchor age group with the standard linear transformation

$$\bar{A} = \left(\frac{S_a}{S_*}\right)X_a + M_* - \left(\frac{S_a}{S_*}\right)M_a$$

where $\bar{A}$ is the estimated anchor-level raw score; $S_*$ the standard deviation of anchor level raw scores; $S_a$ the raw-score standard deviation for age group $a$; $X_a$ a particular raw score in age group $a$; $M_*$ the mean of raw scores for the anchor level; and $M_a$ the mean raw score for age group $a$.

3. Instead of preparing norms separately for each of the 18 age groups, an aggregate distribution of estimated anchor-level raw scores for all 626 cases was developed, resulting in the derivation of standard scores for the entire standardization sample.

4. For each domain, frequencies and cumulative percentages of estimated anchor level raw scores were determined for the aggregate distribution. Cumulative percentages were plotted on arithmetic probability paper, and a smoothed curve was drawn through the plotted points.

5. Percentile ranks for each anchor-level raw score were read from the smoothed curve and recorded. Normal deviates ($z$) corresponding to the percentile ranks were read from tables of the normal curve and were transformed to normalized standard scores with a mean of 100 and a standard deviation of 15, using the formula $15z + 100$.

6. After the unsmoothed standard scores corresponding to anchor level raw scores were plotted on arithmetic graph paper and a smoothed curve was fitted through the plotted points, smoothed standard scores were read from the curve and recorded.

7. Raw scores for each of the initial 18 age groups were then converted to smoothed normalized standard scores using a standard linear transformation.

8. Interpolation was used to determine normalized standard scores for the 4 month (for ages 0 months to 1 year 11 months) and for the 6 month (ages 2 years to 7 years and 11 months) age groupings. A steady progression of
standard scores across all age groups was obtained with slight smoothing of
the norms tables.

_Total test composite standard score_: Normalized standard scores with a mean
of 100 and a standard deviation of 15 were derived for the complete test in
accordance with the procedure outlined by Angoff and Robertson (1983). The
steps for the same are given below:

1. For each of the 626 standardization sample participants, domain standard
scores were obtained from the norm tables.

2. For each of the 18 age groups, means and standard deviations of the average
domain standard scores were computed. In addition, a mean and standard
deviation for the total sample of 626 individuals was determined.

3. Average standard scores for each of the 18 age groups were transformed by
application of the following formula to form a single aggregate distribution
\((N=626)\) of average standard scores:

\[
\bar{A}_t = \left( \frac{S_t}{S_a} \right) X_a + M_t - \left( \frac{S_t}{S_a} \right) M_a
\]

where \(\bar{A}_t\) is the estimated average standard score for the composite
distribution; \(S_t\) the standard deviation of average standard scores for the total
sample \((N=626)\); \(S_a\) the standard deviation for average standard scores for age
group \(a\); \(X_a\) a particular average standard score in age group \(a\); \(M_t\) the mean
average standard score for the total sample; and \(M_a\) the mean average standard
scores for age group \(a\).

4. Frequencies and cumulative percentages were prepared for the aggregate
distribution of average standard scores. Cumulative percentages were plotted
on arithmetic probability paper, and a smoothed curve was drawn through the
plotted points.

5. Percentile ranks corresponding to estimated average standard scores were read
from the smoothed curve developed in step 4. Normal deviates \((z)\) from tables
of the normal curve were determined for each percentile rank and then
transformed to unsmoothed standard scores with a mean of 100 and a standard
deviation of 15 according to the formula \(15z + 100\).
The unsmoothed standard scores corresponding to estimated average standard scores were plotted on arithmetic graph paper and a smoothed curve was fitted to the plotted points. Smoothed standard scores were read from the curve and recorded to develop a master table of smoothed standard scores for the total test composite score.

**Percentile ranks:** Percentile ranks which correspond to the normalized standard scores for the various domains of the test and total test composite score were developed by using the tables of the normal curve.

**Classification categories:** The descriptive categories for the five domains and the total test composite, were determined according to the number of standard deviations the standard scores fell above or below the mean. The classification categories used included: very superior, superior, high average, average, low average, borderline, mild retardation, moderate retardation, severe retardation, and profound retardation.

**Description of the Constructed Battery**

The final constructed battery has 378 items which measure performance of the child in five developmental domains. These domains further have two sub-domains each, with items ranging from 21 to 73 items, designed to assess the development and functioning of children from birth to 7 years and 11 months. Norms for the interpretation of the Development Battery included standard scores at two levels: domain level and overall level. The standard scores ranged from 20 to 160; i.e., from 5½ standard deviations below the mean to 4 standard deviations above the mean. For the total DQ, the values of the total sample approximated the desired mean of 100 and standard deviation of 15 (M =100.10, SD= 14.90). The Social-Emotional DQ mean was 102.25 (SD= 14.14); Adaptive Behavior DQ mean was 102.47 (SD =11.22); Motor mean DQ was 105.04 (SD =10.49); mean Communication DQ was 98.69 (SD= 10.76); and mean Cognitive DQ was 103.71 (SD=10.44).

The entire battery takes about 60 to 75 minutes to complete and this is similar to the time taken by other comprehensive development batteries.

**Reliability**

Three types of reliability estimates were calculated: Split half or internal consistency reliability, test-retest reliability, and interrater reliability. Two statistics
were used to report reliability, namely the reliability coefficient and the standard error of measurement (SEM).

**Split-half or Internal Consistency Reliability**

Internal consistency was determined by calculating the split-half reliability coefficients obtained for the five domains and for the total score separately for all the 18 age groups. Split-half correlations for the sub-domains were computed using raw scores for the two halves of items for sub-domains and these correlations were then corrected using the Spearman-Brown formula.

Examination of the Development Battery reliability coefficients revealed that total score reliability coefficients ranged from 0.90 to 0.97 across the 18 age groups with a median of 0.94. These coefficients were highly acceptable. However, the domain coefficients were less acceptable and the median coefficients were 0.63 for the Social-Emotional Domain, 0.78 for the Adaptive domain, 0.79 for the Motor domain, 0.79 for the Communication Domain, and 0.89 for the Cognitive Domain. The median correlations were highest for the Cognitive domain and lowest for the Social-Emotional domain. The split-half reliability coefficients for the total composite score were excellent and more satisfactory than the domain split half estimates.

**Test-retest Reliability**

During the standardization phase of the study, the test was administered twice to 36 children and their parents. The children were spread across the age ranges from 9 months to 87 months. The battery was re-administered after a median gap of 12 days. The results indicated that the test-retest reliability coefficients for the domains and the total composite score were very high. For example, the coefficients ranged from 0.93 for the Communication domain to a high of 0.99 for the Social-Emotional domain. The test-retest reliability coefficient for the total test composite score was 0.95 which is excellent. In general, the average differences between scores from the first testing and the second testing were small. The means reported for domains and total test score were generally slightly higher for the second administration.

**Inter-rater Reliability**

Objective scoring procedures were defined for each item, and the items administered to the child in a single session were re-scored by a trained examiner to
create a second set of item scores to compare with the original examiner’s item scores. Both examiners calculated domain standard scores and total test composite scores based on the responses they received from the caregiver and their interpretation of the child’s responses to the test battery. These two sets of scores obtained from the two examiners were then correlated to obtain the inter-rater reliability. The inter-rater reliability was very high and ranged from 0.94 (Communication domain) to 0.97 (Cognitive domain). The mean difference in the standard scores was low and ranged from 0.67 (Cognitive domain) to 1.77 (Communication domain).

To conclude, the internal consistency reliability coefficients for the total score were very high across all the age groups and were highly acceptable. However, the domain coefficients were less acceptable and lower than the accepted standard of 0.80 to 0.90. The test-retest reliability coefficient for the total test composite score was also high and excellent and the mean differences between scores from the first testing and the second testing were small. The inter rater reliability was also above 0.90.

**Standard Errors Of Measurement**

The standard errors of measurement (SEM) were computed for the five domains and the total test composite using the standard deviation of 15 in conjunction with the split half reliability coefficients. In addition, the SEM for subdomains were calculated by using the raw score standard deviations and the split-half reliability coefficients obtained for each of the 18 age groups. The results indicated that the SEM values ranged from 2.10 to 4.65 standard score units (Median = 3.60) for the total composite score, with most of the SEMs ranging from 3 to 4 standard score points. The domain specific SEMs were, however, much higher. Thus, the total DQ was a more reliable measure of the child’s true score than the domain DQs.

**Confidence Intervals**

The standard scores were banded with confidence intervals that take into account errors of measurement. Bands of error were calculated for five levels of confidence: 68%, 85%, 90%, 95%, and 99% and are reported for the battery at the domain and the total DQ level. The values for the 68% confidence correspond to the standard errors of measurement. To obtain values for the other four confidence levels, the standard errors of measurement were multiplied by standard deviations.
corresponding to known areas under different portions of the normal curve. The multiplier for the 85% confidence level was 1.44, for the 90% confidence level was 1.64, for the 95% confidence level was 1.96, and for the 99% confidence level was 2.58.

Validity

Three types of validity were established for the battery: content, construct, and criterion related validity.

Content Validity

The content validity of the battery was assessed by examining three dimensions: professional judgment of content, coverage of important constructs, and empirical item analysis conducted at the time of the pilot phase. Coverage of important constructs was also determined by experts who examined whether each of the five domains was addressed by the items selected in the initial pool. Item analysis was conducted only in the tryout phase and items with a poor fit were deleted.

Construct Validity

Construct validity was established by examining the developmental progression of scores; intercorrelations of domain and total test scores; and factor analyses.

The construction of the Development Battery was based on several general developmental principles. First, a child who performs well in one developmental domain should perform well on other domains as well. An examination of the intercorrelations of the domain standard score and Total DQ scores for the entire standardization sample revealed that the highest correlations between the domains were between the Cognitive and Motor domain (0.50) and between Communication and the Cognitive Domain (0.48). The lowest correlation between the domains was between the Motor and the Social-Emotional Domain (0.28). The correlations generally were the highest between the domain and the Total DQ score for all the domains. The correlations between the domains were in the low to moderate range, since each domain represented one unique aspect of the child’s development. The low to moderate correlations between domains also indicate that development was interrelated and all the domains were to some extent related to one another.
Second, developmental abilities should increase with age from birth to about eight years with the maturing of the brain. Since the Development Battery was designed to measure developmental growth, differences in the scores across age groups provided evidence of construct validity. The results indicated that the raw scores of the various sub-domains and domains consistently increased with the age of the child. The domain curves across age ranges showed that before 3 years of age, the Motor, Adaptive, and Social-Emotional curves showed the steepest slope or more growth and then showed a flattening out at older ages. The Cognitive and Communication domains, on the other hand, showed relatively a more flat slope in the initial years and then a steady steep slope till about 8 years with no signs of flattening. This demonstrated that the motor, adaptive, and social-emotional development is the most rapid in the earlier years while the cognitive and communication development coincided with the age when the child began attending pre-school and continued developing till 8 years. The positive growth trends for all the developmental domains, with more rapid development at younger ages, provides evidence for construct validity of the Development Battery.

Factor analyses were employed to gain information for the validity of the structure of the Development Battery. The principal component analysis was conducted with the domain standard scores to confirm the underlying structure of the test and to determine the percentage of variance accounted for by the first principal component. For each age group, the analysis produced one significant factor, using the "eigenvalue equal to or greater than 1.0" criterion, which for the eight age groups accounted for 47.3 to 64.7% of the variance in domain standard scores. Further, principal factor analysis was conducted for the sub-domain raw scores for each of the eight age groups. Intercorrelation matrices of the sub-domain raw scores, with the effects of chronological age removed, were analysed with the resulting factors rotated orthogonally. It is noteworthy that at the higher age groups, the factor structure became clearer, i.e., the sub-domains fell more clearly in the five domains and a more defined structure emerged.

**Criterion-related Validity**

Criterion validity was determined by examining the correlations of the Development Battery with several existing developmental batteries, intelligence tests,
and test which measure specific domains. These included Developmental Profile II (DP-II; Alpem et al., 1986), Developmental Profile 3 (DP 3; Alpem, 2007), Malin’s Intelligence Scale for Indian Children (MISIC; Malin, 1969), Draw-a-Man test (Phatak, 1986), Bender Visual Motor Gestalt test (Bender, 1946), the Coloured Progressive Matrices (CPM; Raven, 1965), the Vineland Social Maturity Scale (VSMS; Malin, 1971), and the Clinical Linguistic Auditory Milestone Scale (CLAMS; Capute et al., 1986).

Correlations between the domain scores and the DP II (Alpem et al., 1986) were found to be moderate to high, ranging from 0.43 to 0.74 across all domains. The correlations were the highest between the domains assessing similar content. This illustrated the utility of the present Development Battery in providing meaningful and valid information about the child’s functioning in each of the five areas of development.

Correlations between the Development Battery and the DP 3 (Alpem, 2007) were found to be moderate to high, ranging from 0.39 to 0.73 across all domains. The correlations were the highest between the domains assessing similar content. The correlation was the highest (r = 0.82) between the General Development Score on the DP-3 and the total test composite score of the Development Battery.

The correlations between the domain standard scores of the Development Battery and the three IQ scores of the Malin’s Intelligence Scale for Indian Children (MISIC; Malin, 1969) ranged from a low of 0.10 (between Social-Emotional DQ and PIQ) to a high of 0.75 (between Cognitive DQ and FSIQ). Since, the Malin’s Intelligence Scale for Indian Children (MISIC; Malin, 1969) is a measure of the child’s intelligence one would expect that it would be most closely related to the Cognitive DQ of the battery and in this context, it is noteworthy, the Cognitive DQ was highly correlated with the FSIQ (r = 0.83). In addition, the inter-correlations between the total test composite standard scores and PIQ, VIQ, and FSIQ were high and were 0.72, 0.79, and 0.83, respectively. The inter-correlations between the Malin’s Intelligence Scale for Indian Children (MISIC; Malin, 1969) and the Development Battery reflected both convergent validity, by the high correlation of 0.83 between the composite test score of the Development Battery and the FSIQ, and divergent validity, by the relatively low correlations of 0.16 between the Social-
Emotional DQ of the Development Battery and the FSIQ.

The Development Battery was also correlated with the Draw-a-Man test (Phatak, 1993). Majority of the correlations were found to be in the moderate range (0.32 to 0.66). The highest correlation was with the Cognitive DQ \( (r = 0.60) \) and the lowest with the Communication DQ \( (r = 0.32) \).

The validity of the Development Battery was also determined by intercorrelating it to the Bender Visual Motor Gestalt test (Bender, 1946). Intercorrelations between the domain standard scores and Bender Gestalt test ranged from low to moderate with the lowest being 0.27 (Social-Emotional DQ and Bender Visual Motor Gestalt Test DQ) to a high of 0.63 (Cognitive DQ and Bender Visual Motor Gestalt Test DQ). The highest correlation was with the total test composite score \( (r = 0.68) \). The large range in the correlations was to be expected because the Development Battery assesses development in five different domains, whereas the Bender Gestalt test assesses only perceptual motor skills.

The inter-correlation of the Development Battery with the Coloured Progressive Matrices (CPM; Raven, 1965) test ranged from low to moderate with the domain DQs, with the lowest being 0.37 (Social-Emotional DQ). The highest correlation of the CPM was with the total test composite score \( (r = 0.67) \). The range in the correlations was not unexpected because the Development Battery assesses development in five different domains whereas the Coloured Progressive Matrices (CPM; Raven, 1965) assesses only non-verbal abstract reasoning.

In addition to comparing the Development Battery with other tests of intelligence and general development, the battery was also compared to other tests that examine areas which are specific to only one of the scales of the Development Battery. The results indicated that the inter-correlations between the domain standard scores and the Vineland Social Maturity Scale (VSMS; Malin, 1971) ranged from 0.68 (Motor DQ and SQ) to a high of 0.81 (total test DQ and SQ). Interestingly, the correlations were high and similar for all the domains, and not only to the Adaptive DQ as one would have expected, thereby pointing to the construct of general child development that appears to underlie the different areas of functioning. The correlations between the Development Battery and the Clinical Linguistic Auditory Milestone Scale (CLAMS; Capute et al., 1986) ranged from a low to moderate range,
the lowest correlation was with the Motor DQ \( r = 0.26 \) and the highest was with the Communication DQ of the Development Battery \( r = 0.61 \) as one would expect. These results provide strong evidence for the convergent and divergent validity of the present battery in very young children.

**Validity of the Battery for Children Belonging to Different Groups**

Several moderator variables in the standardization sample were evaluated to determine whether the constructed Development Battery could be effectively used across groups without bias towards any group. Comparisons of the groups on each of the five domains and the total development score for the entire standardization sample for five moderator variables including sex, residence, education of the father, education of the mother, and socio-economic status was done. The analyses were conducted to compare the standard scores for each of the five domains standard scores and the total test composite score to the average score for the entire standardization sample, which was 100. Effect sizes were reported to determine whether the statistically significant difference held any clinical meaning.

Overall, the standard differences found between groups were small and did not suggest that they should be interpreted in different ways for different groups or that different norms were needed for children belonging to different groups. This finding is notable in view of the criticisms against many intelligence tests that they are biased against disadvantaged children. Therefore, the present battery seems appropriate for use with children belonging to different marginalized groups.

**Principal Uses of the Battery**

The Development Battery can be used effectively in a variety of settings and for a variety of purposes and can accomplish several assessment and educational objectives as it provides a wide range of norm-based standard scores (20 to 160). Firstly, it can serve both as an assessment tool for assessing global developmental delay in young children and mental retardation in children above five years. Secondly, it can identify both the areas of strengths and weaknesses of the child which can aid in designing an individualized remediation plan for the child. Thirdly, the wide age range facilitates the use of the battery in longitudinal comparisons of the same child over the entire developmental period. Finally, it can also be used as a plan for
intervention by sensitizing parents and school teachers by identifying the developmental tasks in the various domains which the child is deficient in and at which the child needs to work at achieving.

In conclusion, the battery should not be used in isolation while drawing out a remediation plan; as it must always be used in conjunction with other data, such as information derived from history, detailed interviews, observation, and the socio-economic milieu to which the child belongs.

Further Extensions of the Battery

Establishing validity is an ongoing process, and there is, therefore, a need that future research should test the battery with different populations, especially children who are mentally retarded and developmentally delayed. There is, thus, a need to establish discriminant validity and test whether the battery can successfully distinguish typically developing children from children with clinical conditions like autism, cerebral palsy, and language disorders.

There is also a need to assess the use of the battery over a period of time in longitudinal research, both before and after interventions. Longitudinal research needs to be conducted with the battery in order to establish the predictive validity of the battery. Prediction helps in informing the clinician if early alarm or reassurance, based on early assessment, has any basis.

Although the battery provides standardized and percentile scores at the domain and the total test level, it gives no scores at the sub-domain level. Further extensions of the battery should offer some score at the sub-domain level as well. In addition, since standard scores are at times difficult to understand by parents and teachers, age equivalent scores may be generated both at the sub-domain and domain level. Although, age equivalents are problematic scores as they do not compare as well to standard scores in psychometric properties, their addition could help in making the scores more interpretable for the lay person.

Since the primary use of developmental tests is with disabled children, future research and extensions with the battery need to suggest adaptations for children with disabilities. Another welcome addition to the battery would be to construct a short
development screening measure from it for use by professionals, such as pediatricians and teachers, who do not have the time and training to administer a long battery.

In sum, the results support the reliability of the data obtained by the Development Battery and its content, construct, and concurrent validity. The strengths of the Development Battery and the comprehensiveness of the domains it measures all make a strong case of use of the present battery with children as a tool for multi-dimensional assessment, particularly in longitudinal studies, determining developmental trajectories and outcomes, and classifying children.

To conclude, there is a rapidly increasing demand for development screening and assessment instruments. This increase is attributable to the recent screening guidelines from the Indian Academy of Pediatrics, eligibility determination for early intervention, multidisciplinary clinic evaluations, a proliferation of neonatal intensive care unit follow-up programs, and research protocols requiring longitudinal measurement. Developmental screening and assessment is an issue faced by all persons working with children, especially those working with disabled children. Developmental screening and assessment affords early identification of potential problems that can be verified by subsequent evaluation. Developmental outcome, ascertained by assessment, is fast becoming a benchmark to determine the quality of pediatric and child psychology practices. Hence, the development of the present battery with its emphasis on providing a multi-dimensional picture of the child’s current level of functioning is an extremely useful contribution in the field of child development and assessment.