



Consultative Group on Early Childhood Care and Development

MEASURING THE PSYCHOSOCIAL DEVELOPMENT OF YOUNG CHILDREN: THE INNOCENTI TECHNICAL WORKSHOP

by Cassie Landers and Cigdem Kagitcibasi
Summary Report, May 1990

OVERVIEW...	2
CURRENT TRENDS IN EARLY CHILD DEVELOPMENT PROGRAMMES IN THE DEVELOPING WORLD...	3
WORKSHOP GOALS AND STRUCTURE...	6
CRITICAL ISSUES IN THE MEASUREMENT OF EARLY PSYCHOSOCIAL DEVELOPMENT...	7
REVIEW OF EXISTING PSYCHOSOCIAL INSTRUMENTS IN LATIN AMERICA: A FRAMEWORK FOR ANALYSIS...	10
CASE STUDIES...	14
THE DEVELOPMENT AND USE OF PSYCHOSOCIAL INDICATORS: WORLD HEALTH ORGANIZATION...	15
INTERNATIONAL STUDY OF SEVERE CHILDHOOD DISABILITY: THE TEN QUESTIONS SCREENING INTERVIEW...	17
CHILD DEVELOPMENT IN PRIMARY CARE: SRI LANKA...	18
CHILD DEVELOPMENT CHART: INDONESIA...	20
PSYCHOMOTOR DEVELOPMENT ASSESSMENT SCALE (EEDP): CHILE...	21
THE TURKISH EARLY ENRICHMENT PROJECT: TURKEY...	23
SUMMARY...	28
PROJECT OVERVIEW: EVALUATION OF EXISTING EARLY CHILD DEVELOPMENT INSTRUMENTS...	29
REFERENCES...	32
ANNEX C: GLOSSARY OF TERMS...	34
ANNEX D: PSYCHOMETRIC UNDERPINNINGS...	35
ENDNOTES...	37

Overview

From May 7-10, 1990 a workshop was held in Florence to discuss technical issues and practical implications of instruments designed to measure the psychosocial development of infants and young children. The workshop was jointly sponsored by the Consultative Group on Early Childhood Care and Development, New York, USA, and the UNICEF International Child Development Centre, Florence, Italy. This event brought together 24 international child development specialists from 14 countries with expertise in psychology, paediatrics, and programme development.

The workshop evolved from the shared recognition of the unmet need for reliable, valid, culturally relevant measures and instruments designed to describe the early development of children. Relatively high agreement has been achieved regarding indicators and instruments that can be used to measure the physical dimension of development. On the other hand, there is little consensus regarding measures of cognitive, social and emotional dimensions or ways to measure the impact of the childrearing environment on development. The search for appropriate measures is guided by at least four major needs:

- 1. Screening.** In order to target preventive actions there is a need to detect children at risk of debilitated or delayed development. If an instrument is to serve this purpose on a large scale, it must be simple.
- 2. Monitoring.** There does not seem to be an adequate equivalent for psychosocial development of the growth monitoring instruments and measures now in use. Developmental delays can increase with time and therefore it is important to detect the delay as early as possible. Implied are repeated systematic measures to detect changes in the risk of developmental status over time.
- 3. Description/Diagnosis.** There is a need for culturally relevant instruments to identify a child's strengths and weaknesses and to describe supportive and negative features of the child's environment.
- 4. Programme Evaluation.** Without recognized and reliable measures of child development, evaluating the impact of programmes designed to enhance early development is an incomplete exercise, restricted to examining coverage and the logistics of implementation without understanding the impact of the programme on the participants.

In addition to these four needs, a more general goal to be served by existing instruments and measures is the education of parents, community workers and programmers. Instruments can call attention to the need for child development programmes as well as underscoring the benefits of preventive and remedial efforts.

In addressing these needs, it is possible to draw upon a growing body of practical experiences which are well grounded in theory. For example, over the last six years the World Health Organization has been coordinating an exercise to adapt and test instruments in several countries including China, Thailand, India, Pakistan, Argentina and Senegal. The Consultative Group on Early Childhood Care and Development has been addressing this problem by gathering,

synthesizing, and disseminating information on the application of existing instruments. The UNICEF International Child Development Centre in Florence has developed an interest in the measurement area as part of its general concern with the quality of life of the child, child rights, and economic policies affecting children.

Against this background and as a contribution to the need for reliable, valid and culturally relevant child development measures and instruments, an overall project with the following three overall goals is planned:

1. To identify and review existing instruments for measuring the psychosocial development of infants and young children in a select group of developing countries.
2. To define the criteria for choosing among or developing new instruments for a range of programming purposes.
3. To make recommendations regarding the actions needed to strengthen the use of existing measures and to create new indicators in order to enhance the implementation of effective early child development programmes.

In order to realize these goals, the proposed specific project objectives are put forth:

1. To organize an initial Technical Workshop to identify critical issues and to establish agreement on an analytical framework and a set of key questions which will guide the review of existing child development measures and instruments in eight countries.
2. To coordinate and monitor the preparation of country reviews which will be completed during a 9-12 month period. Country reviews will be prepared at the field level under the supervision of a Technical Advisory Panel.
3. To organize a concluding workshop to present country reviews, make recommendations regarding the use of instruments in early child development programmes, and identify areas for further action and research.
4. To prepare a manuscript which would include the country reviews of existing child development measures and instruments as well as a summary of the issues, recommendations and needs for further action.

This Summary Report summarizes the results of the Technical Workshop which was designed to accomplish the first of the four objectives.

Current Trends In Early Child Development Programmes in the Developing World

Before presenting the assumptions underlying the organization and structure of the Technical Workshop, it is important to place the measurement of psychosocial development within the larger early child development (ECD) programming context. This background supports and perhaps adds a note of urgency to the tasks before us.

The neglected subject of child development might well receive greater international and national-level attention as a result of a number of emerging issues and mounting problems:

- The dramatically increasing number of women entering the labour force worldwide, with grossly inadequate responses to the need for developmentally appropriate child care.
- A slowly growing consciousness of the social consequences (school drop-out, juvenile delinquency, teenage pregnancy, drug addiction) of neglecting the young child's development and of the major qualitative improvements needed in primary-level and parental education.
- The severe economic difficulties many societies will face in the coming years if substantial improvements in basic education, starting with education of parents and young children, are not achieved.
- Increasing recognition of the need to ensure the sustainability of recent gains in child survival through a strategy of empowering parents with knowledge about child survival and development; also, recognition of the need for effective and caring childrearing as a key element in the preparation of responsible future parents and mature citizens.

Since the International Year of the Child (1979), hundreds of demonstration programmes and projects have underscored the possibilities for early child development programmes to increase and foster children's abilities to cope with and creatively adapt to their environments. The expansion of child development programmes over the last decade has been dramatic. For example, in Korea during 1982-1986, the percentage of children attending early child development programmes increased from 8 to 57 percent. The expansion of government supported centre-based programmes stemmed in part from the increased need for child care services resulting from the entrance of women into the labour force. This expansion is also seen in Sri Lanka's efforts to cover all 5-year-old children by extending the age of entry into primary school downward and transforming the first year of schooling into a kindergarten for all. Thailand now provides some form of preschool for approximately 24 percent of all children between the ages of 3 and 6 through non-formal centre-based programmes combined with parental education activities. In the Philippines, 19 percent of all children aged 3 to 6 participate in an early childhood enrichment programme which delivers learning opportunities to disadvantaged preschool children in a structured centre-based setting. In China, The All China Women's Federation has created over 200,000 local parent education programmes within the past four years. The curriculum emphasizes the interaction between physical, nutritional and psychosocial dimensions of development. The programme, designed in collaboration with the community, is integrated into existing preschools, primary schools, or in conjunction with periodic medical visits.

While these examples are encouraging, one must not lose sight of the many unresolved issues and complex circumstances surrounding the implementation and sustainability of high quality comprehensive child development programmes for children and families at greatest risk. Although a detailed analysis of the scope and coverage of ECD activities around the world is not available, one is well aware of pockets of inactivity. While considerable progress has been made in many (though not all) countries in Latin America and Asia, programmes are nonexistent in many parts of Africa. Moreover, many innovative programmes find themselves unable to secure sustainable sources of revenue necessary to expand coverage while maintaining programme quality. Often the quality of existing programmes is so poor that positive benefits to the child are minimal, and in some instances negative outcomes may be apparent. Moreover, the current set of

strategies is less successful in reaching high-risk mothers and infants during the prenatal period through the first two years life with integrated programmes that provide the appropriate balance of health, nutrition, and psychosocial components. The vertical thinking and sectoral competition characterizing government bureaucracies and international donors as well as child care providers, often conflicts with this need to provide a comprehensive set of services.

The response to integrated programmes of ECD by the international donor community has been much more difficult to capture. However, the convergence of several major events including The Convention on the Rights of the Child, The World Conference on Education for All, and the World Summit for children indicate a new openness and increased awareness to programmes of early child development. It is hoped that child development specialists will be able to respond to these trends with well-conceived programmes as well as ways to identify, monitor and assess the impact of these programmes on the families and young children they intend to reach.

Workshop Organization

The Consultative Group, in collaboration with the UNICEF Innocenti Centre (ICDC), has attempted to capitalize on the opportunities generated by these trends. In this section the activities of the Consultative Group and the UNICEF International Child Development Centre will be briefly reviewed.

Consultative Group

The Consultative Group on Early Childhood Care and Development began in 1984 as an informal mechanism to facilitate communication and strengthen information related to ECD. The major objectives of this mechanism are to:

- Gather, synthesize, and disseminate information on programme planning and implementation, including such topics as the relationship between health, nutrition and development; education/early intervention; women's work and child care; and policy/programme strategies.
- Advocate among international donor organizations to increase attention on the comprehensive needs of the young child and the range of available programmatic alternatives.
- Provide technical assistance to sponsoring agencies with a particular emphasis on programme evaluation and training.
- Encourage the active exchange of information through the operation of a network comprised of over 1,000 individuals from 120 countries. Members of the network receive a bulletin and supplementary mailings of materials related to early child development.

Over the years, 15 agencies have participated in this informal mechanism including UN agencies, private foundations, bilateral technical assistance agencies and non-governmental organizations (NGO). The largest financial contributors have been provided by UNICEF, The Ford Foundation, and US Aid for International Development (USAID). In addition, the work of the

Consultative Group could not have been accomplished without the technical backstopping provided by the High/Scope Educational Research Foundation located in Ypsilanti, Michigan.

The International Child Development Centre

UNICEF's International Child Development Centre (ICDC) located in Florence, Italy, has been in operation since September 1988. The Centre is housed within the Spedale degli Innocenti, a founding hospital designed by Filippo Brunelleschi in 1445. The Spedale is one of the great architectural works of the Early European Renaissance.

In keeping with the Spedale's 550-year mission for children, the primary purpose of the ICDC is to strengthen the capacity of UNICEF and its cooperating institutions to foster an emerging global ethic that responds to the needs of children and their families. In accomplishing this objective, the Centre provides opportunities for the exchange of information among professionals in various fields and undertakes or promotes policy analysis, applied research, and dissemination of concepts directed towards the goals of child survival, protection, and development. Over the next decade, the ICDC aims to become both a vehicle and a voice for changing public policy related to children.

The initial activities of the ICDC are concentrated in four major programme areas: (1) National Capacity Building for Child Survival and Development in Africa, (2) Economic Policies and Mobilization of Resources for Children, (3) Rights of the Child, and (4) Needs of the Urban Child. The Centre's core programme expenses and basic operating costs are financed by the Italian government.

Since its inception, the Innocenti Centre and the Consultative Group have embarked on several collaborative ventures. Most notably, the First Innocenti Global Seminar on ECD was held in June, 1989 in which 24 senior level programme officers and their counterparts participated in a three-week seminar. Not surprisingly, measurement issues cut across all the major themes discussed at the seminar including, for example, the need for instruments to: identify children at highest risk, enhance parental childrearing skills, increase political attention, and monitor and evaluate programme impact. As a result of this recognition, the Workshop summarized in this Report was proposed.

Workshop Goals and Structure

Several factors contributed to the success of this workshop. In the first instance, the background paper prepared by Lucy Atkin, "Analysis of Instruments Used in Latin America to Measure Psychosocial Development in Children from 0-6 Years," provided a masterful starting point. In that paper, a framework outlining a set of criteria to be applied in the selection and analysis of existing instruments was presented and critiqued.

Second, the workshop capitalized on the insights generated from seven years of experience in the design and implementation of a study across six countries co-ordinated by the World Health

Organization (WHO). This project was designed to; (a) identify simple reliable, and culture-appropriate measures of psychosocial development and physical growth, (b) develop a framework for establishing the environmental factors which influence growth and development, (c) prepare locally acceptable home-based monitoring records which can be used by mothers and paraprofessionals.

Capitalizing on the experience of these initiatives combined with participants' expertise, the overall goal of this initiative is to prepare a set of reviews which will identify, synthesize and critique the existing child development measures in a given country or region according to a well-defined framework. Based on these reviews, recommendations regarding the actions needed to strengthen existing measures and their use will be put forth. As outlined in the Workshop Agenda (Annex B-not available) the four-day workshop included the following:

- Presentation of the framework and its application in Latin America.
- Case study presentations and discussions.
- Small group discussions on the issues raised by the case studies and the development of a revised framework based on these discussions.
- Participant country review proposals.
- Specification of project timeline and preparation of second Technical Workshop proposed for September, 1991.

The following sections of the report summarize the issues and recommendations for further action which evolved from these discussions.

Critical Issues in the Measurement of Early Psychosocial Development

State-of-the-Art Review

The recent shifts from medical to social science models of health, from a focus on treatment to prevention of disease and promotion of health, and from an individualistic perspective to a familial-communal perspective, call for different assessments and indicators of the developmental process. One of the most important forces behind this movement is the belief that health is as much a social-behavioral issue as a medical issue. This long overdue perspective underscores the need for multidisciplinary approaches to the study of health and well-being. These multidisciplinary approaches are relevant to recent developments in policy-oriented research that look beyond issues of survival towards the quality of development for those children who manage to survive. Within this context, multi-disciplinary work carried out by psychologists and anthropologists together with medical scientists and health workers assumes a level of great importance.

These important shifts from child survival to child development and from a biological-medical model to a contextual-interactional model of health require measures of psychosocial aspects of human development that parallel the existing measures of physical growth. The sophisticated and culturally valid conceptualizations of human development must be operationalized if they are to have an impact on programmes. Operationalization of the contextual-interactional model of human development requires comprehensive measures which assess physical growth, sensory-motor, language, cognitive and socio-emotional development. In addition, measures that assess aspects of physical environment, environmental stimulation, and child-care-giver interactions are needed. The demands of field level workers, policy makers, and others involved in providing services and introducing interventions converge with the above requirement. There is a need, therefore, to assess both the capacity of the environments to promote healthy psychosocial development as well as aspects of the child's total development.

There is a common assumption that measuring psychosocial development is too complex or cannot be done. Psychologists have not yet sufficiently challenged this assumption. Some components of existing measures have considerable cross-cultural relevance and could be pilot-tested through existing child growth monitoring endeavors. Unfortunately, the general mistrust of psychological measures has resulted in a tendency to ignore this important sphere of early child development. This has, in turn, further reinforced the application of biological or medical models in conceptualizing issues in health and development.

There are regional variations on the availability and use of child development instruments. Latin America, for example, is more advanced than other areas of the developing world. In general, however, instruments assessing psychosocial development that are based on indigenous experience and research are scarce. Many factors have contributed to this scarcity. In the first instance, the resources and expertise required for test construction and standardization are not often available. Secondly, there is an uncritical acceptance of western instruments by professionals in developing countries who have themselves received training in these Western models. Finally, an individualistic approach which takes the child out of context is still prevalent. In this framework the impact of cultural and contextual factors on development receives little attention.

The application of Western instruments, with little or no modification, often creates a mismatch between the purpose and use of the instrument. Norm-referenced western instruments (See Annex D for a discussion of norm-referenced and criterion-referenced test construction) are often inappropriate yardsticks for measuring children from very different socio-cultural contexts. As a result, children in culturally diverse populations with different child rearing orientations and communication patterns are often inappropriately labeled. The use of these instruments for screening or diagnostic purposes, therefore, must be approached with caution. Furthermore, detection and/or screening efforts must be accompanied by appropriate follow-up and treatment services. Screening is not an end in itself. It is only one step in an integrated process of early detection and screening which must be coupled with appropriate interventions and further diagnostic evaluations.

These factors call for caution in the use of western instruments in non-western contexts and for concerted efforts to construct culturally sensitive and ecologically valid instruments, to assess the psychosocial development of children. In accomplishing this task the extensive accumulated experience in assessment and psychometrics both in the western and the developing world needs to be acknowledged and strengthened.

Although some valid measures of psychosocial development are available, much needs to be done in order to enhance the cross-cultural validation of existing instruments. As well, efforts to fill identified gaps through the development of culturally sensitive instruments must be supported. Ecological-cultural validity is at least as important as the reliability of the measures. This is because an assessment instrument that is valid and reliable in one cultural context may maintain its internal consistency (reliability) but fail to measure the construct that it is supposed to measure in another cultural context. Known group comparisons and other tests of validity need to be conducted to establish an instrument's effectiveness within differing cultural contexts.

In practical terms, field workers require instruments for detection and screening, periodic monitoring, description and diagnosis, and programme monitoring and evaluation. For example, a recent workshop held in Brazil concluded that:

A simple instrument to measure children's psychosocial development needs to be perfected and applied in the different countries of the region. Current methods are considered, by and large, to be inadequate.

Though it is unlikely that any single instrument can be developed, the statement reflects the need for practical and time-efficient measures, which can be easily administered at a low cost. This poses a challenge for psychologists who are not often familiar with the constraints imposed by programmatic concerns.

Existing child development instruments utilize both direct (observation, testing) and indirect (interviewing) data collection methods. Even though the importance of the period from birth through age six is recognized, the majority of instruments focus on the period from birth to age three. By contrast, the period from 3 to 6 years of age, when development of language and communication, problem-solving, socio-emotional and like skills occurs at a dramatic pace, has been less adequately addressed. The dominant role of pediatricians and nutritionists as well as the emphasis placed on physical growth and sensory-motor development have contributed to this imbalance.

From a cross-cultural perspective, measures to assess aspects of the environment known to affect child development are even less well established. The very notion of "environmental risk" must be re-conceptualized and translated into programmatic terms. For example, general indicators such as family socio-economic status or parental education are not adequate and more refined measures are needed. An unraveling of the environment is called for if we are to discover the mediating factors associated with child development. While culture-specific environmental measures need to be developed, basic general categories may be identified cross-culturally. These could include, for example, parental interaction styles, expectations, attitudes and beliefs, cognitively stimulating environments, and physical characteristics of the environment.

The general challenges presented here are further explored in the following section of this chapter. In this section instruments measuring the psychosocial development of children from birth through the first six years of life in use in Latin America are reviewed.

Review of Existing Psychosocial Instruments in Latin America: A Framework for Analysis

Introduction

As highlighted in the preceding section, the challenges confronting both the identification of existing instruments as well as the creation of new techniques to measure the psychosocial development of young children are vast and complex. Much, however, has already been accomplished and the insights and experiences of these efforts need to be captured. For example, the paper by Lucy Atkin reviews the creative efforts throughout Latin America to develop and/or adapt instruments designed to measure the psychosocial development of children from 0 to 6 years of age. The paper proposed a framework and a set of specific criteria for the selection and evaluation of instruments. The analysis was organized by classifying instruments into four main categories including, (1) detection or screening for developmental disabilities, (2) periodic monitoring or assessment of a child's development, (3) description or diagnosis of at-risk children, and (4) programme evaluation.

The following discussion briefly summarizes the framework as well as the recommendations generated by the review of existing instruments in Latin America.

Conceptual Foundations

Underlying each instrument is a particular conceptual framework of child development and the factors which influence it. Although the conceptual foundation is not often explicit nor subject to critical analysis, it should be seriously considered when developing or selecting instruments. The basic principles which underlie Atkin's conceptualization of child development and the factors which influence it are briefly summarized as follows:

- Child development is a continual, on-going process which proceeds through qualitatively different stages in an individually unique and irregular fashion.
- Child development is multifaceted, involving diverse processes which are, however, inter-related and inter-acting. Development includes diverse areas of performance which can be conceptually separated but mutually influence each other to varying degrees.
- Development is a dynamic, interactive process in which both the environment and the child's characteristics reciprocally influence each other. The child's characteristics influence and modify the impact of the environment on development. The way the environment treats a child will be partially determined by the way the child is perceived.

- Certain self-righting tendencies can be observed throughout the developmental process. Many children possess remarkable resiliency in the face of stress and are able to overcome early disadvantages.

Criteria for Selecting and Evaluating Instruments

As indicated in Table 1, Atkin identified seven criteria which can be used to examine the quality of existing instruments as well as to guide the development of new measures. A primary criterion is a clear statement of the instrument's purpose. Lack of attention to the purpose of an instrument, combined with ambitious combination of multiple purposes, has created instruments which do not correspond to any particular purpose. For example, some descriptive instruments are too simple to generate informative data, while detection and screening tools are often too complex. A second criterion is the need for clearly defined *objectives* specifying what aspects of development the instrument is attempting to measure. Third, the selection of behavioral *indicators* should be appropriate for the instruments' objectives and for the population in which the instrument is to be used. A fourth criterion requires that the instruments be *culturally appropriate*. Fifth, the instruments should have adequate concurrent and predictive *validity*. This is an area in which the major limitations and needs are most obvious. Sixth, continued efforts are needed to assure adequate *reliability* and *consistency* among practitioners. A final criteria, in the selection of indicators, relates to the ease with which instruments can be incorporated into *programmes*. Instruments have been developed that have simple attractive formats, easy-to-produce materials, and both practical and simple ways of reporting results. The remarkable and painstaking efforts to develop simple measurement alternatives indicate the feasibility of incorporating such tools into a variety of community-based programmes.

TABLE 1—PSYCHOSOCIAL INSTRUMENTS:

Evaluation Criteria

Purpose
 Objectives
 Indicators
 Culturally Appropriate
 Concurrent and Predictive Validity
 Reliable and Internally Consistent
 Programmatic Interface

This set of criteria was used to describe and evaluate a range of existing instruments available in Latin America. The discussion of specific instruments was organized according to the

instrument's purpose, description/diagnosis, detection and screening, periodic monitoring, and programme evaluation.

The instruments included in the review were:

- Escala de Evaluacion de Desarrollo Psico-Motor (EEDP, Chile)
- EDIN (Costa Rica)
- Escala Brasileira do Desenvolvimento de Heloisa Marinho (EDHM, Brazil)
- Milestones of Child Development During the First Six Years (Instituto de Desarrollo Infantil Colombia)
- Ficha Individual de Observacao do Desenvolvimento (Short form of the EDHM, Brazil)
- Short versions of the EDIN
- An analysis of growth and development charts in use in nine Latin American countries.
- A health card incorporating developmental indicators, developed in Uruguay by the Latin American Center of Perinatology and Human Development
- A 12-question screening interview developed by Gensini and Gavito of the Child Development Institute, Cali, Colombia.

Conclusions and Recommendations

As a result of the analysis of existing instruments in Latin America, a set of conclusions and recommendations which correspond to each of the criteria listed in Table I are described. These conclusions and recommendations are summarized as follows:

Purpose

The purpose of the instrument is often unspecified. The following steps are recommended:

- Careful attention should be paid to specifying the instrument's purpose in relation to the kind of population in which it will be used
- It is preferable that one clear purpose be proposed for a particular instrument and fulfilled adequately rather than trying to accomplish too many things with one instrument
- Greater attention should be given to incorporating carefully selected risk factors and environmental indicators.

Objectives and Appropriateness of Indicators

The behavioral objectives are often unclear. As a result, the behavioral indicators are inappropriate for the task. Even when objectives are clearly defined, practitioners may use the instrument with other objectives in mind. In other instances, instruments are used to

inappropriately identify age-specific difficulties that are beyond the instrument's scope. The following steps are recommended:

- Instruments should clearly specify which aspects of development they are designed to describe and/or detect.
- The developmental areas or problems should be chosen according to the programme's priorities and to the feasibility of describing and/or detecting age-specific developmental difficulties.
- Behavioral indicators not included in the traditional tests should be considered where relevant. Attempts should be made to incorporate indicators derived from recent research findings on resiliency/vulnerability factors within the child and the family environment as well on precursors of later cognitive development.
- More attention should be given to the development of instruments dealing with important environmental processes related to developmental outcome. These should include an understanding of indigenous practices as well as relevant factors which place families under stress.

Culturally Appropriate

In Latin America widespread concern for culturally appropriate instruments has stimulated a remarkable amount of creative innovative work. In most cases this has been translated into the design of original and attractive formats to obtain age appropriate norms for behavioral milestones derived from other tests. Less attention has been given to the development of new indicators. To what extent such "natural indicators" could facilitate the educational impact of the instruments as well as upholding the principals of concurrent and predictive validity remains to be seen and should be the topic of careful analysis. The following steps are recommended:

- More attention should be given to the identification and validation of culturally appropriate indicators and their appropriate incorporation into specific instruments.
- The language methods used to describe behavior should be carefully designed and tested in order to optimize parental and personnel comprehension. Jargon should be avoided wherever possible. Pictures should be used instead of instruction manuals where feasible and appropriate.
- Practical research should be undertaken to evaluate practitioners' and beneficiaries' level of comprehension.

Reliability and Validity

This is the area in which the major limitations and needs are most obvious. Although more attention has been given to inter-observer reliability, continued efforts are needed to assure adequate reliability in the hands of the people who actually are using the instruments. In some cases the internal consistency of individual developmental components has not been sufficiently evaluated. Moreover, little attention has been paid to instruments' concurrent and/or predictive

validity. This is particularly important for detection/screening instruments. The following steps are recommended:

- Reliability and validity studies should be carried out before large scale implementation.
- Modifications should be made according to the results of such studies in order to optimize instruments' reliability and validity.
- A coordinated effort should be made to develop practical models for validation studies which optimize resources.

Programmatic Interface

Although instruments with simple attractive formats, easy-to-produce materials, and simple ways of reporting results have been developed, they have not been evaluated to determine their application and efficiency within a programmatic context. The following steps are recommended:

- Continued efforts should be made to design simple culturally relevant instruments with direct applications in community-based programmes.
- The extent to which the instruments are used and understood should be evaluated through focus groups as well as through systematic observations within the setting.

The analytical framework presented by Atkin provided a starting point for continued discussion. The following section summarizes the case studies which were presented and critiqued by workshop participants according to the criteria outlined in Table 1. Through these discussions, the framework was modified and adapted. As discussed in a later section of this report, the analytical framework which evolved from these discussions will be used to assess the effectiveness of existing instruments in eight of the countries represented at the Workshop

(The reader is encouraged to consult the full document for a detailed analysis of the instruments cited here, as well as other examples which are provided within the review. The document is available on the CD-ROM: Early Childhood Counts: Programming Resources for Early Childhood Care and Development, created by the Consultative Group on ECCD, and also available at <http://www.ecdgroup.com>.)

Case Studies

Presentation of six case studies provided an opportunity for participants to apply the assessment criteria defined by Atkin. In the first instance, the World Health Organization's (WHO) initiative to develop a psychosocial screening instrument through the identification of age-specific, culturally relevant milestones was presented. The second case study reviewed the Ten Questions Interview, a screening instrument used in a number of countries designed specifically to detect children with severe disabilities. The issues involved in the development and implementation of monitoring and assessment tools were discussed through the initiatives in Sri Lanka, Indonesia, and Chile. Finally, the case study from Turkey presented a set of instruments used to evaluate the impact of a maternal intervention programme on both mother

and child outcomes. These case studies did not include instruments designed to diagnose and describe children with developmental delays. While critical to the implementation of interventions for at-risk children, these more sophisticated instruments, used by highly trained professionals, were beyond the scope of the present workshop.

The Development and Use of Psychosocial Indicators: World Health Organization

In 1983 the World Health Organization established a working group to determine the need for a set of psychosocial indicators to complement those used for height and weight. The group concluded that the need for such measures was critical since none of the scales currently available were suitable for universal application. Most of the existing tests were developed in the West, and the items and standards were not necessarily suitable to the social, economic, and cultural contexts of other countries. Other theoretical discussions led to an interest in collecting data on environmental factors known to affect child development. Thus, the need for a framework to collect data about environmental factors associated with growth and development was also acknowledged. Once obtained, this data could be used to establish culture-specific predictors of physical growth and psychosocial development.

In response to this need, a protocol was developed and applied in six countries, including Argentina, China, India, Pakistan, Senegal, and Thailand. The specific objectives of the collaborative study were as follows:

- To develop a process leading to the establishment of locally relevant screening tests (Basic Set of Psychosocial Tests and Measurements of Growth) based on the growth and psychosocial development of children from 0-6 years.
- To establish norms for each population or sub-population studied.
- To devise a home-based screening record to use with the tests.
- To develop a framework to examine the environmental factors which lead to optimal growth and those which put children at-risk.

■ DEVELOPMENT OF SCREENING INSTRUMENTS

In each of the participant countries, screening instruments were developed in two distinct phases. The first phase consisted of instrument development and pilot-testing, including: (a) selection of tests and preparation of instructions for test application and scoring, (b) tester training, (c) reliability checks, and (d) repeated administration and refinement. It also involved the progressive deletion, addition, and modification of specific test items.

The second development phase involved the application of the instrument on a large sample (5,000-10,000 children) which produced culturally specific reference values. This phase also provided an opportunity for conducting a field test of the feasibility of integrating the screening

instrument into existing growth monitoring activities. Similar steps could be applied in the development of the family health interview designed to measure the environmental factors which place children at risk.

It is acknowledged that different cultural patterns exist in different cultures, and a skill important in one setting may be unnecessary in another. Another difference between cultures may be that the same milestone is reached at different ages. Thus, all the items included in the protocol were reviewed and rejected or modified according to local relevance. In other instances, culture-specific items were developed. Therefore, the list of specific test items provided by the WHO protocol provided guidance in the identification and selection of test items.

The items included in the screening instruments were designed to assess several developmental domains, including gross motor, fine motor, vision, hearing, language, concept development, self-help and social skills. Items included in each of these domains were selected according to the following criteria: items should (a) be simple, low cost and easily integrated into existing Primary Health Care Systems, (b) assess behaviors that were recognized and valued by parents and primary health care workers, (c) be easily observed and measured by a simple pass or fail score, and (d) be appropriate for use in the measurement of population health status over time.

■ METHODOLOGICAL CONSIDERATIONS

Given the complexity of test construction, one must be cognisant of the common problems and issues involved in the application of the WHO protocol. For example, questions related to sampling (whether to include the elite and/or marginal groups), stratification, and sample size are complex and required country-specific solutions. In addition, since physical and psychosocial development are age dependent, mechanisms to determine the age of the child must be developed. This again can present enormous difficulties when birth certificates are not available. Thus, creative solutions are required including, for example, the use of local events calendars.

Even the use of standardized testing protocols, however, does not ensure against error, and the importance of inter-rater and test-retest reliability studies must not be underestimated. Precautions against mistakes must be taken repeatedly, and intensive tester training programmes should be carefully designed and monitored. In all the participating countries, quality control aimed at reducing residual error was an essential element.

■ RESULTS

At the present time considerable progress has been made in China, India, Thailand and Pakistan in the development and field testing of screening items. For example, in China 8,995 children under six years of age from 46 areas in the Southeast Provinces of China were tested. In addition, a Chinese growth and development home-based record has been developed in which risk is represented by colors. Three centres in India (Chandigarh, Tabalpur and Hyderabad) have collected data on both psychosocial and growth indicators as well as factors related to the family environment for 14,200 children. In Thailand, data on 5,350 children from 11 different centres has been collected and analyzed using the Grostat analytical package. Finally, Pakistan's research team has collected data on 3,000 children from urban and rural areas. In addition, 1,200 families

have been interviewed with a standardized protocol. The data from both of these protocols are being analyzed.

International Study of Severe Childhood Disability: The Ten Questions Screening Interview

The Ten Questions Screening Instrument (TQ) was designed as part of a pilot study coordinated through Columbia University, New York, USA. The purpose of this instrument is to detect children from three to nine years of age with severe developmental disabilities including mental retardation and psychiatric disorders, cerebral palsy and other motor disorders, blindness, deafness and epilepsy.

The interview, which is administered to the child's caregiver, consists of four questions concerning vision, audition, movement and convulsions, while the remaining six questions focus on the child's cognitive development. The ten questions included in the interview are listed below.

1. Compared with other children, did the child have any serious delay in sitting, standing or walking?
2. Does the child have difficulty seeing, either in the daytime or at night?
3. Does the child appear to have difficulty hearing?
4. When you tell the child to do something, does he/she seem to understand what you are saying?
5. Does the child have difficulty in walking or moving his/her arms or does he/she have weakness and/or stiffness in the arms or legs?
6. Does the child sometimes have fits, become rigid, or lose consciousness?
7. Does the child learn to do things like other children his/her age?
8. Does the child speak at all (can he/she make himself/herself understood in words; can he/she say any recognizable words)?
9. For 3- to 9-year-old children ask: Is the child's speech in any way different from normal (not clear enough to be understood by people other than his/her immediate family)?
For 2-year-old children ask: Can he/she name at least one object (for example, an animal, a toy, a cup, a spoon)?
10. Compared with other children his/her age, does the child appear in any way backward, dull or slow?

If responses to any one of these questions are in the wrong direction, a child is referred for further diagnosis. For example, a child with positive responses to questions 1, 2, 3, 5, 6, 9, 10, and one or more negative responses to questions 4, 7, and 8 is considered at-risk for a serious developmental problem.

A validation study of the instrument has been carried out in nine developing countries. The study design compared the Ten Questions Interview with a more detailed detection questionnaire. Independent evaluations were performed on all positive screenees by local experts as well as on a random sample of negative screenees. The final report of the project (Belmont,

1984) indicated that in the majority of the countries, the Ten Question Interview (TQ) had better sensitivity than a complete interview. Although the target objective of the instrument is to detect severe mental retardation, it was also able to identify children with more subtle developmental problems.

This instrument provides an interesting example of how very simple questions may be used to screen certain populations for specific developmental problems. The pilot study should also be recognized as an instrument validation model which can be used in community-based programmes.

Child Development in Primary Care: Sri Lanka

The development and identification of appropriate child development instruments and screening procedures has taken place within the context of Sri Lanka's primary healthcare programme. Initiated seven years ago, the programme's primary objective was to promote child development through a system of early detection coupled with appropriate intervention. The programme was developed as a result of several factors including (a) the existence of an extensive primary health care system including clinic and home-based monitoring of child health and nutritional supplementation for high-risk groups of pregnant mothers and children (0-5 ages), (b) recognition that parents had little confidence and knowledge of their capacity to improve the quality of their children's experiences, and, (c) recognition that parents and health care workers had limited knowledge about children's developmental needs and the interactional nature of the developmental process.

The selection, monitoring, and refinement of instruments was an integral part of the programme. It was possible to carefully integrate the selection and use of the instruments into the ongoing programme without increasing costs and expanding the programme with new fixed targets. Indicators for monitoring the health and development of children as well as the quality of home environment were critical for programme evaluation as well as the identification of at-risk children requiring specific intervention.

As a result of this process, indicators of the following variables were selected and incorporated into a Child Health and Development Record (CHDR).

1. Weight
2. Developmental milestone
3. Quality of family environment (in relation to defined risk factors)
4. Health compliance
5. Immunization
6. Family planning

The primary objective for integrating the milestones into the existing Child Health Record was to educate parents through primary health care workers. The milestones included in the chart, as well as the variables used to measure the quality of family environment, are listed on the Tables 2 and 3. The milestones for monitoring children's development were identified by selecting items from well-established development norms for children from birth to 5 years of age. The criteria

for the selection of the milestones included feasibility of observing the behavior in the home setting, comprehensibility, and sensitivity of the item to detect children at-risk. In addition to the chart, interviews of primary caregivers were conducted by health workers. The interview procedures are outlined in Figure 1 (figure 1 not available).

TABLE 2—CHILD HEALTH AND DEVELOPMENT RECORD

<i>Development Milestone</i>	<i>Age/Month</i>
Smiles Responsively	1-2
Holds Head Up	3-4
Sits Without Support	8-9
Stands Without Support	9-10
Walks Alone	11-15
Says, "Amma," (That)	10-15
Follows Simple Instructions	15-30

TABLE 3—CHILD HEALTH AND DEVELOPMENT RECORD

Family Environment

- Mother Under 17 Years
- Two Children Under 3 Years of Age
- Mother or Caretaker Dull
- Mother Mentally Ill
- Mother or Caretaker Disinterested in Child
- Disorganized House
- Father Known to be Delinquent /Alcoholic or Otherwise Mentally Disordered
- Severe Marital Discord
- Abject Poverty

A study of the Child Health and Development Record was conducted in urban, semi-urban and rural areas. A total of 600 households were visited by pairs of health workers. Inter-rater and test-retest reliability for observational items reflecting between 68-85%. Correlations at .01 level of significance were found between items reflecting knowledge of child care, household organization, play opportunities, interest in and knowledge about the child, level of poverty, marital discord and child behavior (curiosity, lack of curiosity, irritability, unhappiness, activity,

frequent tantrums). An association was also found (.05 -.01 level of significance) between immunization compliance, child spacing, number of children, knowledge of children's needs and opportunities for play.

In addition to these results, a UNICEF-supported longitudinal study on perinatal and infant mortality and morbidity validated the association between home risk factors and health compliance during pregnancy and the perinatal period. Studies of similar developmental screening activities from India and Indonesia also generated similar findings. When combined with the flow chart, the sensitivity and specificity of the CHDR to detect children at risk was greatest for children between 36-40 months of age. These interesting and promising initial findings underscore the potential of the CHDR as an effective screening and intervention strategy which can be easily integrated into ongoing systems of child health monitoring.

Child Development Chart: Indonesia

In Indonesia child survival goals have been achieved in many regions of the country. Thus, programmes optimizing the development of children's cognitive as well as physical potential are in demand. The goal of this project was to design a child development component which could be incorporated into the Indonesian National Growth Monitoring Programme. This component takes the form of a child development chart that can be used at health posts by volunteer health workers in conjunction with the ongoing administration of the growth chart. The child development chart can be used to support positive caretaking behaviors and to increase care-givers' skills in providing cognitively stimulating environments.

Tufts University School of Nutrition, Massachusetts, USA in collaboration with Diponegoro University School of Medicine, Java, Indonesia, designed an innovative intervention programme to encourage positive indigenous caretaking behavior to improve child growth and development. An innovative and effective tool, this development chart has been adapted from several child development checklists and scales commonly used in the United States and Indonesia. The chart can be used to monitor the development of children between 0-36 months.

The design of the Child Development Chart evolved from the experience of a successful child stimulation project sponsored by the Ford Foundation in Central Java. In that project, 253 children in Mlonggo, a rural region of Central Java, were studied over a six-month period. The social, behavioral and physiological factors that operated to protect children in spite of severely limited household and community resources were investigated. Anthropometric measurements, physical and mental development scores based on the Bayley Scale of Infant Development and levels of stimulation based on the Cadwell Inventory of Home Stimulation (HOME) were collected. The scales used for determining home stimulation were adapted to Javanese culture.

The analysis indicated that the quality of maternal-child interaction was related to positive growth and development outcomes in spite of distressful environments. As measured by the Cadwell Inventory, infants whose mothers understood and responded to their developmental needs were significantly better nourished and had higher mental development scores. The findings also indicated that better nourished children had significantly advanced physical and

cognitive development. The results of this study were used in designing developmental and nutritional messages. In addition, the chart and responding child stimulation activities capitalized on the success of projects in which significant developmental gains had been obtained by combining psychosocial stimulation with nutritional interventions.

Through a series of both qualitative and quantitative techniques including focus group discussions, concept testing, and formative evaluation, the Indonesian chart and messages were created. The practices, and beliefs of mother and kaders (community mothers acting as health volunteers) regarding child stimulation, development, and nutrition were determined through focus group discussions. Through this process a cultural specific inventory of child development milestones was developed. A sample of the developmental milestones for children from birth to 2 years of age and accompanying messages are listed in Table 4. Each developmental milestone is accompanied by a “message” to the mother on how to stimulate her child's development. Good nutritional practices and their important relationship to optimal cognitive and physical development are also emphasized.

At the present time, the chart is being tested, evaluated and modified through a series of five test-cycles. At each cycle, the chart is used to determine the developmental stage of approximately 200 children during the monthly growth monitoring session in 12 posyandus (health posts). At the start of the first cycle a psychologist, assisted by a kader, tested the children to determine the developmental stage and to review the age-appropriate stimulation activities. During the following month, the kader made home visits to ensure that the mothers had understood and were implementing the activity. The testing of children's development and the teaching of new stimulation activities were carried out on a monthly basis. At the end of each two-monthly test-cycle, the chart and the messages were evaluated. These recommendations were incorporated and a modified chart was used in the subsequent test cycle.

The initial results of this rather elaborate testing phase are tremendously encouraging and underscore an exciting potential for the integration of child development monitoring into existing growth monitoring systems.

Psychomotor Development Assessment Scale (EEDP): Chile

The Psychomotor Development Assessment Scale (EEDP) assesses psychomotor development during the first two years of life. The instrument was developed by Centro de Estudios de Desarrollo, y Estimulacion Psicosocial (CEDEP), Roman Diaz 26, Oficina 63, Santiago, Chile, to complement the growth monitoring activities implemented by the National Child Health Programme. It's simpler than a complete descriptive instrument but more complex than most primary screening tools. Children with delays in motor, coordination, social development, or language are referred for further diagnosis and intervention.

The instrument consists of 75 behavioral indicators with five items in each age group (one item from each of the developmental domains). The items, adapted from existing infant assessment tools including the Bayley, Gesell, Brunet and Lezine, and Denver, are easily administered by field workers with little previous training in child development. It requires 15 minutes for

administration and consists of 11 objects which are low-cost and easy to produce. Specific instructions for administering, scoring and interpreting the child's developmental level are described in an accompanying manual. The age at which 25, 50, 75 and 90 percent of a standardized sample of children passed a given item are reported. Raw scores are then converted to standard scores and a developmental quotient is generated. A graph and developmental profile greatly facilitate interpretation of the results.

TABLE 4—THE CHILD DEVELOPMENT CHART: SELECTED ITEMS

<i>Age (Months)</i>	<i>Developmental Milestone</i>	<i>Stimulation Message</i>
0	Inspects surroundings	
1	Smiles in response to person	Hold baby facing you and talk and smile. Play everyday.
2	Holds head up	Place baby on stomach. Talk to him.
3	Eyes follow dangling ring	Dangle item. Let baby watch it move.
4	Makes 3 different sounds	Talk to baby, tell him what you're doing.
5	picks up small item	Sit the baby up, give the baby small items to reach for.
6	Looks for fallen spoon	Hide toys and help the baby find them.
7	Sits alone steadily	Let baby sit up with a little help.
8	Plays peek-a-boo	Play "peek-a-boo" with the baby.
9	Picks up item thumb and one finger	Give the baby small items to pick up.
10	responds to verbal request	Help baby to wave or clap while saying the words.
11	Uncovers toy	Cover toy while baby is looking. Help him find it.
12	Walks along	Let the baby walk holding on to your hands.
13	Imitates adult	Show child how to do what you do.
14	Stacks 2 cubes	Give child small boxes or blocks of wood to stack.
15	Says two or more words	Encourage child to repeat names of objects/activities.
16	Brings objects on request	Ask the child to bring you things. (2-part directions)
17	Points to three body parts	Teach the child parts of the body.
18	Feeds self with spoon	Teach the child to use spoon
19	Speaks in 2 word sentences	Expand 1-word sentences into 2-word sentences.
20	Points to three pictures	Point to pictures in KMS card and name

		them for child.
21	Builds a tower of four cubes	Give child small boxes or blocks of wood to play with.
22	Discriminates 2 objects	Tell your child stories.
23	Kicks ball/balances on 1 foot	Show your child how to kick a ball.
24	Names three objects	Have child repeat names of objects after you.

The EEDP reports adequate reliability for individual items (test-retest) as well as the overall scale. Constructive validity studies have been performed with regards to age and socio-economic status. However, the concurrent and predictive validity of the instrument has not been ascertained. Without such data it is difficult to determine how the EEDP compares and contrasts to other available tests and clinical assessments. In summary, the EEDP is relatively easy to learn, can be applied by intermediate level personnel, and can be easily integrated into practical clinical settings.

The Turkish Early Enrichment Project: Turkey

The Turkish Early Enrichment Project was a four-year (1982-1986) longitudinal study of the developmental impact of two early child enrichment interventions implemented in low-income areas of Istanbul. Two types of programmes, namely center-based and home-based, were evaluated, separately and in combination. Funding for this project was provided by the International Development Research Centre (IDRC) in Canada.

The project involved children in three types of settings, including existing custodial day care, educational day care, and children cared for at home. The impact of these three "context categories" (custodial, educational and home) on the development of 252 three and five-year old children were investigated. In years two and three of the Project a home intervention or mother training programme was administered to about half of the mothers of randomly selected children in each of the three groups. The research design was a three (context: custodial, educational, home care) by two (age: 3 or 5) by two (maternal training or no maternal training) factorial design.

■ PROCEDURE

In the first year, baseline data was collected on the children's cognitive and social-emotional development through behavioral testing and observation. Maternal interaction and maternal teaching style were studied through the observation of a structured task. Maternal interviews were also administered to assess maternal child rearing orientations, life-styles, self-concepts and satisfaction with life. These interviews also provided information about the home environment as well as other demographic data. In the second and third years of the project, a home-based

maternal training intervention was delivered to randomly selected groups of mothers with children in custodial, educational, or home care. In the fourth year post-tests were carried out to determine differences between the experimental (trained) and control (non-trained) groups. In order to understand and promote the overall development of the child in several different domains combined with a holistic contextual framework, a wide range of measurement instruments were required (Table 5). In addition, multiple administrations were used in order to increase validity. Almost all of the baseline or pre-test assessments used in the first year were repeated in the fourth year except for those which were inappropriate for older children (7 and 9 years of age). Additionally, extensive school-related assessments were administered in the fourth year of the project including school achievement and attitudes toward school.

■ MATERNAL HOME-BASED INTERVENTION

The maternal intervention programme included two components: "cognitive training" and "maternal enrichment". The cognitive training was based on the HIPPY programme (Home Instruction Programme for Preschool Youngsters) developed in Israel. This extensive learning package for mothers of children between 4-5 years old was implemented over two 30-week periods. The programme was conducted through a network of para-professional field workers drawn from women within the community. Mothers who were trained at home or in group settings were instructed to systematically apply cognitively stimulating techniques with their children. In this way, mothers assumed the role of trainer or teacher of their own children.

The "maternal enrichment" component of the home-based intervention was implemented through group discussions held on alternate weeks. The topics covered ranged from health and nutrition, psychosocial needs of the growing child, communicating with the child, discipline, and expressing feelings. Enabling mothers to better cope with problems and attend to their children's needs as well as their own was a primary objective of this intervention.

■ RESULTS

The two main independent variables were context of development (custodial day care, educational day care, and home care) and presence or absence of mother training. While the results on the context of development will not be presented here, one finding deserves mention; namely, educational day care was superior to custodial day care and home care for virtually all indicators of psychosocial development. This was particularly significant in regard to cognitive development and school achievement. The results indicate that early child development was enhanced only through day care programmes that provided cognitive stimulation within an educational milieu.

The analysis of the maternal training found positive effects on children's overall development and school achievement (Table 6). Children whose mothers had training surpassed the control group of children on measures of cognitive and social development. In none of the measures did the control group excel. It appears that even without directly providing the children with early stimulation, parental education programmes can exert a powerful impact on children's development.

As summarized in Table 7, the impact of the intervention programme on maternal behaviors and attitudes also found significant results. In terms of their interactive styles with their children the findings show greater responsiveness, higher levels of verbalization, higher achievement aspirations and expectations from their children. All of these undoubtedly contribute to the superior level of development demonstrated by the child outcomes in Table 6.

Direct effects on mothers were indicated by a higher intra-family status vis-a-vis the spouse and greater optimism and outlook on their lives.

In summarizing these findings, it appears that this maternal-focused intervention enhanced the developmental outcome of the children by promoting maternal well-being. Thus, this approach is able to address the intersecting needs of mothers and children. It should be recognized that two thirds of the mothers were factory workers with minimal education levels. This flexible parental education intervention has the potential for widespread application. Ways to integrate this model into existing health and nutrition education programmes should be explored.

TABLE 5—CHILD, MATERNAL, AND FAMILY INDICATORS

CHILD DEVELOPMENT INDICATORS

Cognition

IQ Scores:	The Stanford-Binet; Goodenough Draw-a-Person test.
Cognitive Development:	Piaget Seriation and Classification tasks.
Complexity of Behavior:	Tizard six-level rating scale of observations.
Cognitive Style and Analytical Thinking:	Children's Embedded Figures Test and WIPPSY *Analytical Triad"

Personality and Social Development

Autonomy-Dependence:	Observation rating and Mother Interview
Aggression:	Observation rating and Mother Interview
Delay of Gratification:	Mischel procedure
Self-Concept/Self-Esteem:	Draw-a-Person; Larsen and Leigh Index; Mother Interview
Emotional Indicators:	Draw-a-Person
Level of Social Participation:	Tizard six-level rating scale of observations
School Adjustment/Liking:	Mother Interview

Maternal Indicators

Child Rearing Attitudes and Behaviors:	Mother Interview
Life Style of the Mother:	Initial and Mother Interview
Self-Concept of the Mother:	Mother Interview
Mother's Satisfaction with her Environment:	Mother Interview
Mother's Teaching Style:	Hess and Shipman Toy and Block Sorting Task

Family Indicators

Background and Demographic:	Initial Interview
Home Environment:	Mother Interview
Physical Environment Index	
Environmental Stimulation Index	

TABLE 6

SUMMARY OF RESULTS OF MOTHER TRAINING: CHILD OUTCOMES

Cognitive

Stanford Binet IQ:	Statistically Significant
Draw-a-Person:	Trend in the expected direction
WIPPSY Analytical Triad:	Statistically Significant
Block Design:	Statistically Significant
Children's Embedded Figures:	Trend in the expected direction
Piagetian Classification:	Statistically Significant
Piagetian satiation:	Not Statistically Significant
 Achievement Tests	
Math:	Trend in expected direction; Statistically Significant at younger age group
Turkish:	Trend In expected direction
General Ability:	Statistically Significant
School Records:	Statistically Significant (overall average); Satisfically significant In Turkish and near statistically significant insocial studies and math (increasing effect)
School Adjustment	Near statistically significant trend In the expected direction

Personality/Social

Dependence:	Near statistically significant trend In the expected direction
Asking for Help:	Statistically Significant
Aggression:	Statistically Significant
Self-Concept:	Near statistically significant trend in the expected direction (Rohner)

TABLE 7

SUMMARY OF RESULTS OF MOTHER TRAINING: MATERNAL OUTCOMES

Interaction with Child

Full Attention to Child:	Statistically Significant
Interaction when Together:	Statistically Significant
Read or Tell Stories to Child:	Statistically Significant
Teaching Cognitive Skills:	Trend in the Expected Direction
Help with Child's Homework:	Statistically Significant

Attitudes/Expectations

More pleased with Child Now:	Statistically Significant
Child Needs Love:	Statistically Significant
'Good Child" Concept:	
More Industrious, Autonomous; Less Obedient Expect Child (to do):	
School Success; Less Self-Care and Household Tasks Expect Child to Succeed in School:	Trend in the Expected Direction
Desire more years of schooling for child:	Statistically Significant
Expect more years of schooling for child:	Statistically Significant

Values /Discipline/Style

Autonomy as Pleasing Child Behavior:	Statistically Significant
Verbal Communication when Pleased:	Statistically Significant
More Reasoning and Advising for	

Misdeed:	Statistically Significant
Less Physical and Verbal "punishment":	Statistically Significant
More Intentionality, More Praise, Positive Feedback and Cognitive/Rational Appeals In Teaching Style:	Statistically Significant

Direct Effect on Mothers

Shared Decision Making:	Statistically Significant
Communication:	Statistically Significant
Mother Decides Child Discipline:	Statistically Significant
Shared Activity:	Statistically Significant

Outlook on Life

Better off Now than Before:	Statistically Significant
Due to Education:	Statistically Significant
Optimism toward Future:	Statistically Significant

Summary

Several conclusions can be drawn from the discussions stimulated by these innovative attempts to measure the development of young children. As emphasized through the case studies, the purpose for which the instrument has been designed (screening, monitoring/assessing, diagnosis, or programme evaluation) must be clearly specified and understood by those administering the instrument. Test results must be assiduously interpreted so as to avoid problems of misclassification. In particular, the information presented to parents must be carefully tailored to meet their concerns and prevent misinterpretation. As indicated by the studies, one must be cautioned against the use of instruments with over ambitious or multiple purposes. The wide range of needs of practitioners and policy makers underscores our need for multiple instruments and systems of detection, monitoring, and diagnosis combined with adequate follow-up and treatment. Until such systems are in place, the selection of instruments should be based on their sensitivity to detect specific problems as well as the availability of appropriate referral and intervention programmes.

Several of the presentations illustrated the value of incorporating environmental indicators which describe processes known to affect developmental outcome. The discussions highlighted the importance of capturing environmental processes which increase or decrease a family's vulnerability to stress and subsequent risk of child rearing deficiencies. The lessons of existing design and analysis protocols to this effect will provide an important foundation for further activity.

All of the instruments reviewed responded to the demand for practical, attractive, easily administered instruments which can be integrated into existing growth and nutrition monitoring services. In this regard, the initiatives in Indonesia, Chile, and Sri Lanka were particularly sensitive to the need to adjust and modify items from existing tests in order to identify cultural specific age-appropriate norms. Participants agreed, however, that efforts to create new items which capture the unique set of skills in a given setting should be enhanced.

Much of the discussion focused on the rather complex psychometric issues of test construction including for example, reliability, validity, sensitivity, and specificity. The need for a concentrated effort to meet fundamental psychometric requirements was underscored. In spite of these demands, one was comforted by the existing knowledge base and availability of individuals with expertise in this area. Thus, it is the responsibility of those embarking on new initiatives to learn from and build upon the lessons of the past. As illustrated by the case studies, practicality, reliability and validity protocols have been developed. The challenge of confronting workshop participants is to optimize resources through the application of existing test construction and validation protocols within and between countries.

Finally, and perhaps most importantly, was the emphasis placed on the need to appraise the existing and/or potential use of instruments within the programmes they support. For example, instruments could be more effectively used in the design or planning of early child development programmes through the identification and description of target populations as well as the specification of objectives and programme goals. As was seen in several instances, the instruments can play a critical role in the education of communities, paraprofessionals, families, and parents about issues of early child development. Finally, as illustrated by the Turkish Case study, instruments, carefully matched to programme objectives, can be used to monitor the process and evaluate the impact of programmes on both parents and children. Convincing arguments in support of early child development programmes can be made on the basis of these results.

Thus, while the challenges loom large, the discussion motivated participants' commitment to fulfill Workshop objectives. The following section summarizes the specific project proposal to review existing instruments in eight countries as well as the suggested actions for future research.

Project Overview: Evaluation of Existing Early Child Development Instruments

Framework Overview and Application

As indicated, the primary goal of the Workshop was to develop a revised framework for the analysis of existing instruments as well as to generate a set of guidelines for the selection of instruments to be included in the eight proposed country reviews. The revised framework is portrayed in Table 8. As a result of workshop discussions, a fundamental modification in the framework was the need to specify and strengthen the link between an instrument's psychometric characteristics and its use within the programming process. Thus, in addition to a critique of the

instrument according to the criteria outlined in the Table, the instruments' potential or current application in the design and evaluation of ECD programmes will be assessed. The column headings appearing in the Table are summarized below while definitions of the specific items listed under each heading can be found in Annex C: Glossary of Terms and Annex D: Psychometric Underpinnings.

Instrument Assessment Criteria

- 1. Background.** For each instrument list the name, author, date of development and address of where the instrument can be obtained. A copy of each instrument should be included.
- 2. Measurement Purpose.** Describe the purpose of the instruments, including: monitoring/assessment, screening/detection, description/diagnosis.
- 3. Characteristics.** Determine whether the test measures aspects of the child and/or attributes of the child's immediate environment. Also indicate the age range for which the test is appropriate, its format and the level of expertise required to administer the instrument. Provide data on the time required for administration and how often the test needs to be applied in order to obtain meaningful results.
- 4. Developmental Domain/Environmental Attributes.** For each child development instrument, describe which and how the various developmental domains are measured including: motor, sensory, cognition, language, emotional, social/life skills. For instruments that measure aspects of the child's immediate environment, describe what and how they are measured including, physical environment, family environment, social support systems, and patterns of care-giver/child interaction. (Please note: Attributes of the child's environment need not be limited to those items indicated on the Table.)
- 5. Psychometric Properties.** Each instrument selected should be critiqued according to the psychometric characteristics identified in the Table including reliability, validity, sensitivity, specificity, standardization, norm and criterion-referenced. In addition, the appropriateness of the instrument for the culture in which it is being applied should be discussed. Please refer to Annexes C and D for definitions of terms listed under this heading.

Use within Programmes: Actual/Proposed

The purpose of this assessment is to evaluate the effectiveness of the instrument in the design, implementation and evaluation of early child development activities. If the test has not been widely used, its potential applications should be examined.

- 1. Background.** Describe the context(s) or service settings in which the instrument is currently used. Include, when available the number of settings using the instrument.
- 2. Programme Planning.** Describe how the instrument is or could be used to identify programme goals and objectives and to identify the size and scope of the target population in need of services. Describe how the instrument is or could be used to gather data useful in the planning of implementation strategies or in the design of materials and methods. The instrument's feasibility including material and human costs and time required for administration should also be assessed.
- 3. Programme Implementation.** Describe how the instrument is or could be used in delivering early child development programmes. Specifically, is the instrument useful in

educating parents, field workers, and/or the community? Describe how the instrument is or could be used directly to improve the child's development. Describe how the instrument is or could be integrated into an overall assessment of the child, including health and nutritional status.

4. Programme Evaluation. Describe how the instrument is or could be applied in the ongoing monitoring of child development programmes as well as its ability to measure the short or long term impact of the programmes on participants.

5. Additional Comments. If necessary provide additional information regarding any aspect of the instrument or its application within the programming process. This section could describe personal observations or anecdotal information as, for example, staff's, parents' and children's reactions to the instrument, cultural sensitivity, and ease of administration.

Selection of Instruments

The selection of instruments for review is a critical step. Since the primary audience is composed of policy makers and practitioners, the selection should be biased towards the inclusion of simple and practically oriented instruments. Thus, highly detailed and technically oriented instruments designed for detailed diagnostic and/or academic research purposes should not be included in the review.

Following the Second Technical Workshop a working group could be established to identify sites for pilot testing, design reliability and validity studies, and determine strategies for the development of procedures for large scale field testing within existing community-based programmes. This process could contribute to the development of culturally sensitive and valid instruments.

Another goal of the Second Technical Workshop is to identify gaps in available instruments and propose strategies for constructing new instruments. The prospects for constructing new instruments may vary along two basic dimensions including (a) the targeted age range (0 to 3 or 3 to 6), and (b) child development or environmental indicator. A particularly exciting challenge, therefore, would be to address the need to develop environmental indicators known to affect the development of the child from 3 to 6 years of age. The urgent challenge of this task is to determine whether cross-culturally stable variables or categories could be established in the face of great diversity of child rearing values, behaviors, and environmental conditions.

In summary, the overriding challenge and opportunity confronting workshop participants is to construct culturally sensitive and valid instruments in order to evaluate the rapid growth of early child development programmes in the developing world. It is hoped that the issues summarized in this Report combined with the review of instruments in eight countries and suggested proposals for future activity, underscore the willingness of the Workshop participants to accept these challenges.

References

- Atkin, L. (1989). "Analysis of Instruments Used in Latin America to Measure Psychosocial Development in Children from 0 to 6 Years of Age." A paper prepared for the Consultative Group on Early Childhood Care and Development, UNICEF House, New York, New York.
- Belmont, L. (1986). "Screening for Severe Mental Retardation in Developing Countries: The International Pilot Study of Severe Childhood Disability." In Beg, J.M. (Ed.). *Science and Service in Mental Retardation*. London: Methuen.
- Belmont, L. (1984). *The International Pilot Study of Severe Childhood Disability, Final Report: Screening for Severe Mental Retardation in Developing Countries*. Utrecht, The Netherlands: Bishop Bekkers Foundation and Institute Research Series, No. 1.
- De Silva, M., Nikapota, A., Vidyasagara (1988). "Advocacy and Opportunity: Planning for Child Mental Health in Sri Lanka". *Health Policy and Planning*; 3 (4): 302-307.
- Durkin, M., Zaman, S., Thornburn, M., Davidson, L "Population-Based Studies of Childhood Disability in Less Developed Countries: Rationale and Study Design". In press, *International Journal of Mental Health*, 1990.
- Kagitcibasi, C.; Sunar, D.; Bekman, S. (1989). *Comprehensive Preschool Education Project: Final Report*. Ottawa: IDRC.
- Kagitcibasi, C. (1990). *The Turkish Early Enrichment Project*. UNESCO, Notes, Comments ..." (forthcoming).
- Landers, C. (1990). "Innocenti Global Seminar Early Child Development Summary Report, June 12-30." UNICEF International Child Development Centre, Florence, Italy.
- McGuire, J. and Austin, J. (1987). "Beyond Survival: Children's Growth for National Development." *Assignment Children*, 2, 1-52.
- Myers, R.G. and Hertenberg, R. (1987) "The Eleven Who Survive: Toward a Re-examination of Early Childhood Development Program Options and Costs." *The World Bank Report*, NO. EDT 69.
- Myers, R.G. (1989). Consultative Group on Early Childhood Care and Development, (1989). Report of a Workshop on Early Child Development: Instruments and Measures, Brasilia, Brazil.
- Nikapota, A.D. "Case Study for Sri Lanka -Child Development in Primary Care." Paper prepared for UNICEF, Sri Lanka and Ministry of Health, Sri Lanka. Presented to Innocenti Technical Workshop on Psychosocial Development.
- Rodriquez, S., Arancibia, V., Undurruga, C, *Escala de Evaluacion del Desarrollo Biccomotor 0-24 meses*. Santiago de Chile: Ed. Galdoc, 1979.
- Sockalingam, S.; Zeitlin, M.; Satoto; Coletta, N. (1990). Study to Encourage Positive Indigenous Caretaking Behavior in Improving Child Nutrition and Health. A paper available from the Consultative Group on Early Child Care and Development, New York, NY.
- World Health Organization (1986). Protocols for the Development and Field Testing of Techniques for Monitoring Physical Growth and Psychosocial Development. WHO/MCH/MNH/86.1. WHO, Geneva, Switzerland.

World Health Organization. Progress Report on The Activities of Physical Growth and Psychosocial Development (September 1988-April 1990). Programme of Maternal and Child Health, WHO, Geneva, Switzerland.

ANNEX C: GLOSSARY OF TERMS

Assessment. Ongoing procedures used by appropriate qualified personnel throughout the period of a child's eligibility to identify (i) the child's unique needs, (ii) the family's strengths and needs related to development of the child, and (iii) the nature and extent of early intervention services that are needed by the child and the child's family.

Criterion-Referenced Test. A test that measures a specific level of performance or a specific degree of mastery.

Diagnostic Evaluation. An examination used to ascertain conclusively whether a child has special needs, to determine the nature of the child's problems, and to suggest the cause of the problems and possible remediation strategies.

Norms. Statistics that describe the test performance of specified groups, such as children of various ages or handicapping conditions in the standardization sample of a test.

Reliability. The extent to which a test is consistent in measuring whatever it measures; dependability, stability, relative freedom from errors of measurement.

Screening. A brief assessment procedure designed to identify children who should receive more intensive diagnosis or assessment. Screening is designed to help children who are at risk for health and developmental problems, handicapping conditions, and/or school failure to receive ameliorative intervention services as early as possible.

Sensitivity. A statistical property of a test that indicates the proportion of those children who are at risk who are correctly identified.

Specificity. The proportion of those not at risk who are correctly excluded from further assessment or treatment

Standardized Test. A systematic sample of performance obtained under prescribed conditions, scored according to definite rules, and capable of evaluation by reference to normative information.

Validity. The overall degree of justification for interpreting and using a test's findings. It concerns a test's accuracy. Different kinds of validity evidence are appropriate for different kinds of tests.

ANNEX D: PSYCHOMETRIC UNDERPINNINGS

To understand the basis of the claim that instruments can help improve decisions about children, basic knowledge of underlying psychometric principles is necessary. Several of these principles will be briefly mentioned here. More detailed information can be obtained from Cronbach, L. J. (1984) and Anastasi (1988).¹

Standardization. An instrument consists of a fixed set of stimuli or items and there is a prescribed response procedure. In addition, the instrument is administered under standard conditions described in the instrument manual. Standardization of the instrument and the instrument procedure for different children is a necessary condition if the user wants to compare the results of those children.

Reliability. If a measurement is accurate and precise, it has a small error, and approximately the same score will be found if the measurement is repeated (i.e. if it is repeated under the same circumstances). Accurate measurements with small errors are called reliable in the instrument assessment area. The evaluation of reliability retesting is not the only procedure. Since instruments usually consist of a number of independent items, it is also possible to compare the performance on different parts of the instrument (e.g. odd versus even numbered items). For an instrument with high reliability, individual differences in scores will be consistent across instrument parts.

Validity. A psychometric instrument is usually intended to make inferences about future behavior. The validity of an instrument is a measure indicating the extent to which a certain inference is justifiable and meaningful. More simply, validity shows whether an instrument in fact measures what it is supposed to measure. For example, an instrument can be used to make predictions about a child's future performance in school. Some measure of school achievement (e.g. grade average) then forms a criterion in terms of which the instrument can be evaluated. If it turns out that the instrument score makes it possible to give a fairly precise prediction of the performance on the criterion, the instrument is said to have a good **criterion-related validity** or **predictive validity**. The extent to which a measurement correlates with a criterion when both are obtained at approximately the same time, the instrument is said to have concurrent validity. The main validity issue for detection and screening instruments is their ability to really identify cases (children and/or families) "at risk." Valid instruments correctly differentiate "at risk" cases from normal cases.

We can also ask people with expertise in a certain field whether the contents of the items in an instrument are representative of a certain behavior domain; the **content-validity** is then examined. Experienced teachers are often asked whether an educational test gives a good coverage of the curriculum. In an arithmetic test with items on addition, multiplication and division, items on subtraction are absent. The content validity of the instrument would be improved if such items were added. One can also look at validity from a theoretical perspective. Often it is supposed that an instrument measures a certain psychological characteristic or function, such as being "at risk" or "school maturity". These characteristics are not directly

observable, they are theoretical concepts or constructs. It is possible to collect empirical evidence with respect to the question whether or not an instrument can be considered as a measure of a particular construct. This is called **construct validation** research.

Test Norms. Instrument scores are usually expressed in arbitrary scale units. Consequently, we can do little with the information that a child obtained a score of, let us say, 10 points, indicating that 10 items were done correctly. If we also happen to know that only 5% of the children with the same age as the testee obtained a score of 10 or less, the score can have some relative interpretation. For properly researched instruments information on the level of scores is available in the form of norm tables. These tables are based on an investigation involving a random sample of the population for which the instrument is to be used. In many instances, like developmental scales, a sample is needed of children at various age levels, so that for young children per month and for somewhat older children per half year or per year, a norm table can be constructed.² Thus, norms are data about the performance of groups of instrument takers which are representative of some larger population. Children of a certain age in a given country or region constitute a population in this sense. The tables can be used to assess the standing of an individual testee in the population he or she belongs to. In such cases we speak about **norm-referenced** interpretation of instrument scores.

Sometimes one finds a **criterion-referenced** interpretation of instrument scores. This means that scores are evaluated with reference to some external standard, independent of the performance of instrument takers. A teacher who marks a school test frequently decides in advance which mark will correspond with a certain number of incorrect answers on that test, independent of the actual performance of the pupils in a class. Thus, there are pre-specified performance standards. In criterion-referenced testing an important problem is how to set the criterion. Is a child delayed when he or she cannot walk at an age of 18 months, but is able of most other things which the average 18 month old can do? If an instrument is used for screening purposes, one wants to set the criterion at the lowest level that is not in itself indicative of progressive delay. It is also important to consider which domains of behavior and physical development have to be represented in an assessment instrument, and which items are representative of any given domain once this has been selected for inclusion.³

The distinction between **norm-referenced** and **criterion-referenced** testing is of particular significance for development scales. Whether or not a certain score is indicative of developmental delay can only be properly evaluated against the performance of other children of a similar age growing up in a similar socio-cultural environment as the testee. This points to the need for norm tables, based on a representative sample of children in the population concerned.

However, one can also impose external standards for the behavioral development of children, provided the dangers of culture inappropriateness are clearly recognized. For example, an item could be included asking the 5-year old child to count up to 5 or up to 16. In a society with a low level of literacy most children will fail this item and it will not serve to discriminate children at any age level. At the same time, mastery of this task by children can be seen as facilitating primary school education. In societies where most children cannot count an item as mentioned

can function as a standard for caregivers and community workers which will lead them to teach children this task.

If items are used which are meant to impose desirable standards of behavior, care takers and community workers should be given the opportunity to teach children the relevant skills. This implies that the developmental scale concerned does not only have a screening or diagnostic goal, but that it also acquires an educational function. This is only meaningful if instruction about the content of developmental scales can be included in community development programmes in which caregivers are made aware of the demands that will be imposed on their children, first in the testing situation, but later also in school or in the society at large.

Endnotes

¹ Cronbach, L.J. (1984). *Fundamentals of Psychological Testing* (4th Ed.) New York: Harper and Row. Anastasi, A. (1988). *Psychological Testing* (6th Ed.) New York: MacMillan.

² For intermediate ages, results can be interpreted by interpolation on the basis of the two nearest sets of norms.

³ Confer the notion of content validity.

Early Childhood Counts: Programming Resources for Early Childhood Care and Development.
CD-ROM. The Consultative Group on ECCD. Washington D.C.: World Bank, 1999.