Using Science to Inform Preschool Assessment

A summary report of the
Temple University Forum on Preschool Assessment

January 30-31, 2003

Center for Improving Resources in Children's Lives (CIRCL)

Temple University

Using Science to Inform Preschool Assessment

A summary report of the
Temple University Forum on Preschool Assessment

January 30-31, 2003

Written by
Anita Kochanoff, Ph.D., Visiting Assistant Professor, Department of Psychology
Kathy Hirsh-Pasek, Ph.D., Professor, Department of Psychology
Nora Newcombe, Ph.D., Professor, Department of Psychology
Marsha Weinraub, Ph.D., Laura H. Carnell Professor of Psychology

CIRCL
Translating research into reality

Center for Improving Resources in Children's Lives Temple University

The forum and this report have been made possible by generous support from Temple University's Department of Psychology, Center for Public Policy, College of Liberal Arts, and Office of the Vice President for Research and Graduate Studies.

THE TEMPLE FORUM

The cofounders of the Center for Improving Resources in Children's Lives (CIRCL) at Temple University held a forum on January 30-31, 2003, in order to address what recent research in developmental science can contribute to the best possible implementation of two recent trends in education:

1. An emphasis on early childhood learning, and
2. A concern about accountability.

We begin in Part One with background information for the meeting and the report. Part Two then reviews constructs and measurements in important areas of preschool functioning and provides recommendations. Part Three concludes by describing some of the necessary future directions for preschool assessment.

TABLE OF CONTENTS

EXECUTIVE SUMMARY …5

PART ONE:
Setting the Stage for the Assessment of America's Youngest Children …8
- The Need for Assessment
- The Purpose of Assessment
- Forum Goals
- Challenging the Notion of Milestones

PART TWO:
Constructs and Measurement …10
- Language and Literacy Skills …10
  - Important Constructs to Assess
  - Assessment Instruments
- Mathematical Skills …16
  - Important Constructs to Assess
  - Assessment Instruments
- Social-Emotional Skills …20
  - Important Constructs to Assess
  - Assessment Instruments

PART THREE:
Future Directions in Preschool Assessment …24
- The Need for New Research
- Overcoming Test Biases
- The Need for Ecological Validity
- Constructs Most Relevant for School Readiness
- Integrative Assessment Methodologies
- Sampling Strategies for Accountability
- Conclusions and Action Steps

REFERENCES …28

APPENDIX Forum Agenda …31

FORUM PARTICIPANTS

Hosts:
Kathy Hirsh-Pasek, Ph.D.
Nora Newcombe, Ph.D.
Marsha Weinraub, Ph.D.
Anita Kochanoff, Ph.D.

Language and Literacy Skills

Sue Bredekamp, Ph.D.	CDA Council; Council for Professional Recognition
Jill deVilliers, Ph.D.	Smith College
Roberta Golinkoff, Ph.D.	University of Delaware
Kathy Hirsh-Pasek, Ph.D.	CIRCL, Temple University
Aquiles Iglesias, Ph.D.	Temple University
Susan Landry, Ph.D.	University of Texas, Houston
Fred Morrison, Ph.D.	University of Michigan

Mathematical Skills

Douglas Clements, Ph.D.	University of Buffalo
Herbert Ginsburg, Ph.D.	Columbia University
Janellen Huttenlocher, Ph.D.	University of Chicago
Susan Levine, Ph.D.	University of Chicago
Kelly Mix, Ph.D.	Indiana University
Nora Newcombe, Ph.D.	CIRCL, Temple University
David Uttal, Ph.D.	Northwestern University

Social-Emotional Skills

Clancy Blair, Ph.D.	Penn State University
Susan Campbell, Ph.D.	University of Pittsburgh
Michelle DeKlyen, Ph.D.	Princeton University
Mark Greenberg, Ph.D.	Penn State University
Nathan Fox, Ph.D.	University of Maryland
Anita Kochanoff, Ph.D.	CIRCL, Temple University
Ronald Taylor, Ph.D.	Temple University
Marsha Weinraub, Ph.D.	CIRCL, Temple University

We would like to express our sincere appreciation to the scholars who participated in the Temple Forum for sharing their expertise and for contributing to the writing and editing of this report.

EXECUTIVE SUMMARY

The cofounders of the Center for Improving Resources in Children's Lives (CIRCL) at Temple University held a forum on January 30-31, 2003, in order to address what recent research in developmental science can contribute to the best possible implementation of two recent trends in education:

1. an emphasis on early childhood learning, and
2. a concern about accountability.

Forum Goals

The specific goals for the Temple Forum were:

1. to delineate the key achievements in a sequence of development between birth and 5 years of age,
2. to evaluate what assessment tools currently exist for these milestones, and
3. to determine the extent to which new tools need to be developed or old ones improved.

Experts were convened to examine areas that are relevant to school readiness, specifically: (1) language and literacy, (2) mathematical and spatial skill, and (3) socio-emotional development.

Constructs Most Relevant for School Readiness

Scholars at the Temple Forum reached consensus about important language and literacy constructs that are important to assess in preschool-age children: Vocabulary diversity, Narrative use and understanding, Rhyming, Syntactic complexity, Alliteration, and Quantification (e.g., "each and every, one and only"). Beyond these, scientific evidence shows that to be prepared for school, children also need to be developing numeracy skills, strong problem-solving and other cognitive skills, and social competency.

Recommendations for Existing Assessment Instruments

It is imperative that the assessment tools chosen for program evaluation and accountability produce scientifically reliable and valid data, examine developmental growth, and lead to program improvement. The recommendations resulting from the forum for each of the developmental domains are as follows.

- Language assessments:
  1. The auditory test of the Preschool Language Scale-Fourth Edition (PLS-4) (Zimmerman, Steiner, & Pond, 2002)
  2. Expressive One-Word Picture Vocabulary Test (EOWPVT) (Brownell, 2000)

- Literacy assessments:
  1. Test of Early Reading Ability (TERA-3) (Reid, Hresko, & Hammel, 2001) for print concepts
  2. Auditory processing subtest of the Developing Skills Checklist for phonological awareness (CTB) (McGraw Hill, 1990)
  3. DIBELS (Good, 2000) for letter knowledge

- Mathematics assessments:
  1. Subtests of the Primary Test of Cognitive Skills (PTCS) (Huttenlocher & Levine, 1990)
  2. Test of Early Mathematics Ability-Third Ed. (TEMA-3) (Ginsburg & Baroody, 2003)

- Social-emotional assessments:
  1. Social Competence/Behavior Evaluation Short Form (SCBE-30) (LaFreniere & Dumas, 1995)
  2. Devereux Early Childhood Assessment (DECA) (LeBuffe & Naglieri, 1999)

Need for Development of Improved Assessment Instruments

The recommendations made here are based on the constraint of using available measures until better measures are produced. It is important to recognize that the tools which are currently available suffer from serious problems with respect to their validity for special populations and the over-identification of problems within disadvantaged populations. Researchers at the Temple Forum contend that it is imperative for long-range planning to take place, and funding made available, for the validation of more culturally neutral assessments, some of which are already in development. It is also imperative that new assessments be developed that are more ecologically valid and that help teachers identify those areas in which students need to grow and improve.

The field needs new measures with sufficient breadth and depth of coverage while remaining feasible in terms of time, cost, training and implementation on a large-scale.

We urge funders to seriously consider investing in instrument development; otherwise, current and future studies are caught in a continued cycle: They need appropriate measures but do not have the time or money to develop them. An instrument development initiative would help put an end to this problem.

The development of new assessment instruments will need to address ecological validity, integrative assessment methodologies, test biases, and special populations of children.

Ecological Validity

Most assessments lack ecological validity, rendering them ineffective for informing teaching. They therefore have limited benefit for improving children's learning. Assessments that are process oriented would help teachers learn what is important and might prevent them from "teaching to the test."

Integrative Assessment Methodologies

There are very few integrative assessment procedures which allow for a comprehensive understanding of the nature of children's learning, and that adequately evaluate how competencies in different developmental domains interact to produce optimal functioning and school readiness. For example, a child's progress toward reading and math proficiency in preschool depends on the ability to regulate attention and to use language flexibly in the service of multiple goals. Similarly, progress in social skills and social competence requires the regulation of emotion and the development of a sense of self as an efficacious and active learner. Ultimately the integration of social and academic competence in the regulation of attention and emotion can serve as the basis for ongoing achievement in school. Thus, an integrated assessment of cognitive and social development is important if the field is to progress.

Test Bias & Special Populations

Existing assessments suffer from serious questions regarding their validity for special populations. Evidence suggests that current assessments lead to over-identification of problems within disadvantaged populations, and in children from non-mainstream backgrounds who use other dialectic variants of English or who are English Language Learners. Most of the current tests deal with cultural variation by meeting the criterion of "inclusion," which demands that the standardization sample match the Census data. Minority representation in the standardization sample does not address the possibility that minority children may not perform as well as majority children because of test bias. In the areas of IQ and language tests, it is well-established that these biases exist. Furthermore, when tests mainly assess how children perform to a mainstream standard they draw our attention to deficiencies with respect to that standard, and draw our attention away from equally important information about their proficiencies.

Sampling for Accountability

When preschoolers are assessed for the purpose of program accountability rather than informing practice directly, it is important to develop an effective sampling strategy. A random sampling process, for example, in which a subset of children in classrooms is examined, would produce more in-depth information and more valid knowledge in the amount of time that will be available for testing children due to the limited duration that is appropriate for children of this age.

Also, matrix sampling is an efficient and effective way to collect valid data regarding accountability. Matrix sampling involves giving parts of tests to all children, rather than giving every child all of the test items. This sampling strategy would lower the risk of teachers teaching to the test since it would be impossible to know on which particular items any one child would be tested.

Communicating the Temple Forum Recommendations

As a collaboration of developmental scientists who are invested in informing policy about young children, the participants felt strongly about communicating the consensus reached at the Temple Forum. The forum took place when the federal Head Start program was beginning to develop a standardized assessment tool which would be used to assess four-year-olds several times during their last year in the program. Thus, in addition to the plans for distributing a report, a letter was immediately written to the chair of the advisory group working on the development of this new Head Start assessment, which urged them to consider the recommendations for existing measures resulting from the discussions at the Temple Forum. In addition, a letter was sent to federal funding agencies, expressing the need to support research attempting to establish the validity of more culturally neutral assessments and the development of new assessments that are more ecologically valid.

The Temple Forum was an initial step in addressing the problematic field of preschool assessment with only a small group of the scholars who conduct research on this important topic. The forum was successful in bringing pressing issues to the table and providing initial recommendations. More evidence needs to be examined regarding all the possible measures that may be appropriate to use with preschoolers, and more information is needed about instruments currently undergoing development. The resources collected through this collaboration will serve to refine the recommendations for the best existing assessments and needs for new assessments.

PART ONE: Setting the Stage for the Assessment of America's Youngest Children

The Need for Assessment

Increasing numbers of children are in some type of formal educational setting before entering kindergarten. Recognizing these trends and the need to support families with young children, 41 states devote funding to preschool programming. However, in recent constrained economic conditions, states are faced with difficult decisions about which programs to continue to fund. Thus, policymakers need to know whether publicly funded early education programs are actually benefiting young children.

The proliferation of research on the importance of early brain development and the effectiveness of early interventions with young children has undoubtedly played a role in states prioritizing funding for early education, as well as parents' decisions about their children's early learning experiences. As more and more children are starting the education process earlier, expectations for kindergarten have increased compared to past decades. The discrepancy between children who have had preschool experiences before they reach kindergarten and those who have not seem to be growing larger, making it more difficult for kindergarten teachers to teach the same curriculum to all children in a classroom. Addressing such classroom difficulties is one reason that more schools are conducting school readiness assessments in order to determine which children are prepared enough for the kindergarten experience. The 2002 federal *No Child Left Behind Act*, which aims to "improve overall student performance and close the achievement gap between rich and poor students," is another reason for the rise in preschool assessment nationwide.

The need to demonstrate school accountability, to evaluate programs, and to assess individual children have certainly become issues of national relevance. For example, children in the federal Head Start program will for the first time be given a national, standardized assessment during the year before entering kindergarten. There are considerations, however, that are unique to the assessment of preschoolers, and that may not lend themselves to reliable program evaluation in the way that test scores of older children might.

The Purpose of Assessment

The assessment of preschoolers could potentially serve many purposes. The National Education Goals Panel (Shepard, Kagan, & Wutz, 1998) outlined four primary reasons for assessment, which were discussed at the Temple Forum:

1. Assessment to support children's learning and development by informing teachers
2. Assessment for identification of special needs

3. Assessment for monitoring trends and evaluating programs
4. Assessment for school accountability

It is important to note that some believe accountability leads to high-stakes testing which has the potential to influence important policy decisions such as what programs get funded, which should be closed, and what types of programs should be developed. Scores can also lead to decisions about students and teachers including placement in special education, retention, or advancement.

Forum Goals

The participants at the Temple Forum were convened to address the issues related to the uniqueness of measuring preschoolers' skills and development.

The specific goals for the Temple Forum were

1. to delineate the key achievements in a sequence of development between birth and 5 years of age,
2. to evaluate what assessment tools currently exist for these milestones, and
3. to determine the extent to which new tools need to be developed or old ones improved.

The forum was designed to examine areas that are relevant to school readiness, specifically: (1) language and literacy, (2) mathematics (including spatial ability), and (3) socio-emotional development. To this end, part of the forum was spent in three small working subgroups, in which participants shared their specific expertise for each developmental domain.

Challenging the Notion of Milestones

Before the participants were divided into their respective developmental domain subgroups, there was a discussion about longstanding concepts and terminology that are used in the field of developmental research. For example, the notion of milestones was seen as "a bad legacy" from developmental science because of its past focus on maturation rather than environment. More recent research that emphasizes the role of environment challenges the notion of milestones. Forum participants felt strongly that the interaction between children's development and their environments needs to be examined instead of developmental milestones. Considering learning processes rather than child outcomes provides more useful information, such as children's progress over time and developmental trajectories. The field is now striving to understand individual differences and thresholds in behavior more than purely normative behavior. Assessing processes requires the understanding of the child's context, within which researchers and teachers could consider both the strengths and weaknesses of the child. In terms of what children know or are capable of doing, participants decided that "emerging competencies" was a more appropriate concept than "milestones."

In Part Two, forum participants evaluated assessments using solid, reliable, and valid scientific research and practice to answer the question:

Which existing measures are best suited for the assessment of preschool age children according to the current state of scientific knowledge?

PART TWO:
Constructs and Measurement

Language and Literacy Skills

Members of this subgroup decided that they would discuss measures that met the following criteria.

1. Does the assessment show growth within the specified age range (3-5 years)?
2. In cases of no growth, would the assessment indicate the need for intervention?
3. Is there empirical evidence to support that these are important and achievable expectancies?
4. Is it, or can the assessment be, standardized without regard to culture?
5. Are these constructs considered predictive of school readiness?

In addition, the subgroup operationalized "expectancies" in terms of how they might be charted and in terms of process evaluations instead of product/outcome evaluations. Below is the resulting definition.

- **Expectancies** - "Process dependent abilities that are challenging but achievable across multiple domains for most children within a given age."

Important Constructs to Assess

The participants then delineated the aspects of language development that they believe to be most linked to school readiness. The language constructs that were discussed included words, phonology, syntax, semantics, and pragmatics. Elements of literacy were also outlined.

Words

The subgroup decided that the "expectancies" that would be important at three years of age were lexical organization comprehension, fast mapping comprehension, and word diversity. At age four, the subgroup outlined expectancies for a hierarchical organization of words, growing word diversity, and the use of quantifiers and connectors, morphological compounds, and mental state verbs.

Lexical organization at age three includes contrasts, parallels, and categorization. That is, children are beginning to organize words into related sets. Waxman and Hatch (1992) show that children as young as three, but especially by age four, are beginning to form hierarchies of contrasts for noun classes, such as "This is not an animal, it's a plant," and for the same stimulus "It's not a rose, it's a dandelion." However, finding culturally un-biased categories and tasks may be challenging because these categories (animals, food, clothing, etc.) are more or less enriched by the experiences children have (e.g., with zoos, or with a varied diet or clothing options). It was suggested that there should be more attention to verbs than nouns for a more culturally neutral assessment. Measuring verbs is less culturally biased than current measures, which focus mostly on nouns, because verbs may vary less with respect to the instruction that is received by parents. Similar to the noun study, children can be prompted with a picture for words that are in contrast

[e.g., "He's not walking, he's …" (e.g., crawling) "He's not entering, he's …?" (e.g., going out)]. At age four, the most important change in the lexicon may be the increasing *hierarchical organization* of words and categories, allowing efficient and flexible retrieval (Anglin, 1970, Stockman & Vaughn-Cooke, 1984; 1986; Waxman & Hatch, 1992).

Fast mapping, or the language learning ability, is important and is an ability which has been found in children as young as 12 months with nouns and 24 mouths with verbs. Fast mapping is a characteristic that refers to the fact that children only need minimal exposure to a word to append it to an object, action, or event. It reflects children's current knowledge and their ability to learn new information easily. The subgroup discussed the idea of measuring such abilities using video technology to present new lexical items and scenes (Rice, Buhr & Nemeth, 1990). Example items in determining a child's mapping include: "Give me the chromium one; not the green one, the chromium one" (nouns), or "He zanned the girl; where is the girl zanning the boy?" (verbs). Discovering that a novel word applies to an object or event that is, as of yet, unnamed contributes to more rapid vocabulary growth. Susan Carey (e.g., 1978; 1982), Mabel Rice (e.g., Rice, Buhr, & Nemeth, 1990), and Lila Gleitman (1990) are some of the prominent researchers in the area of fast mapping nouns and verbs (also Pena, Iglesias & Lidz, 2001).

For a measure of vocabulary, the subgroup discussed using a task of word diversity instead of the more common simple word checklist. *Word diversity* refers to the number of different words used by a child; it is an expressive test of "how much of the world the child has mapped and labeled." For example, assessment items may ask the child to list all of the foods s/he can think of. The importance of *word diversity* is demonstrated by Tablors, Roach and Snow's (2002) research in which they found the density of rare words used and understood to be the most predictive factor in word learning.

Word diversity is more culturally sensitive than word lists for many reasons. First, cultures differ in the extent to which they place importance on providing labels for objects or events. Assessments that allow the child to more naturally express items for which s/he has acquired labels may circumvent cultural biases. Second, word diversity can be conducted in any language. Food, clothing, and animals were three topic areas considered as the least culturally sensitive. Furthermore, topics would need to vary enough so that teachers would be less inclined to teach to the test.

At age four, there should be development within lexical categories, such as *modifiers*. In addition, four-year-olds should begin to understand, and effectively use, *quantifiers* (e.g., each, every) and *connectors* (e.g., and, but). For example, children should be able to understand the difference between "He went to the store *and* left the box" and "He went to the store *but* left the box." Most languages have a way to mark such contrasts.

The understanding of derivational *morphology* is also important and emerges around four years of age. Morphology is the addition of suffixes and prefixes, which change the meanings of words. For example, understanding that adding "–er" to a verb (e.g., walk) makes a new noun (e.g., walker) helps in vocabulary comprehension and expression. Research by Eve Clark (1993) provides examples of the types of competencies that should be expected during the preschool years.

Furthermore, four-year-olds should display growth in their comprehension and expression of *mental state verbs*, such as "think," "know," "feel," and "imagine" (Bartsch & Wellman, 1995; de Villiers & de Villiers, 2000; Shatz, Wellman & Silber, 1983). Children should be able to distinguish between another person's belief about the world and the real world. This competency is also important to social-emotional growth and would therefore be a good construct to measure by integrated assessment methods (i.e., multiple features of school readiness are tested in one task).

Mental state verbs are important for school readiness, essential for the "landscape of consciousness" characteristic of genuine story comprehension (Bruner, 1986), and are used in all languages.

Phonology

Phonology, an area of language development currently assessed in Head Start classrooms, is also important for school readiness. The understanding of sounds in relation to language is important in word comprehension and essential to early reading abilities. The participants agreed that there are already good batteries of tests available in this area. The argument was made that because these skills – metalinguistic awareness of sounds, and phonics – are important and helpful in learning to speak and read English, they should be taught in preschool classrooms.

Rhyming is a phonological competency which children typically begin to display around three years of age. For example, children should be able to fill in blanks in songs or raps that have a rhyming scheme. Although rhyming is not common across all languages, for example Spanish, it is important in most English-speaking classrooms. Another important phonological task is alliteration. *Alliteration* refers to the use of the same sound at the beginning of multiple words, for example "Ben is bringing bananas!" Although some tests have shown that alliteration at four years of age is predictive of reading ability, the subgroup knew of no evidence of predictability for children as young as three. Bradley and Bryant's research can speak to this issue.

Additional skills that should be present around age four include *syllable segmentation*, *blending* and *on-set rhyme*. An example of syllable segmentation is the ability to clap out the number of syllables in a word. Blending, which may not be equally common across all cultures, is demonstrated by combining "base" and "ball" to create "baseball." "Tanner, panner, canner" is an example of an on-set rhyme.

Syntax

Many of the important aspects of syntax development are already assessed in preschool age children, and many good tests of these abilities currently exist. However, the subgroup felt that *mean length utterance (MLU)* was not an effective measure of syntax production under common preschool testing conditions. MLU is problematic because its definition varies (e.g., number of words versus number of morphemes such that teach+er would be counted as two). Furthermore, accurate and reliable measures of MLU would require transcriptions which would be unrealistic to record and analyze for large numbers of children. MLU is not comparable across dialects such as African American English in which morphemes such as the past tense are optional.

At age four, MLU loses its utility because the length of utterance varies more with the situation than with the child's competence (Brown, 1973). A higher level version of the mean length utterance competency, which seems to be relatively sensitive to later language abilities, is displaying a diversity of sentence structures. A test which measures this, the IPSYN (Scarborough, 1990) counts the number of types of sentences the child is able to correctly produce, of increasing complexity. However, the feasibility of this task was questioned in terms of being scorable in a classroom setting since it requires considerable linguistic skill, beyond the level of most teachers.

Comprehension and appropriate use of *Wh- questions* was also deemed an important aspect of pre-school language ability. Wh- questions tap the child's syntactic understanding in a sensitive way. At the most minimum level, children should be able to understand the difference between "Who is Bradley?", "Where is Bradley?" and "What is Bradley?" However, a question such as "who bought what?" also taps a child's ability to comprehend how sets can be paired up: Jane bought milk, Sally bought a peach, and so forth (Roeper & de Villiers, 1993). Finally, wh-questions are sensitive to the structure of longer sentences, so the child's answer can reveal what syntax is developed (de Villiers, 1996). Thus, by age four, children should be able to comprehend more

complex syntactical features, such as embedded clauses. This would mean understanding questions such as, "What did the mom say she bought?" compared to "What did mom buy?" (de Villiers & de Villiers, 2000). Three-year-olds would not be expected to display this competency.

In English, *word order* is imperative to sentence comprehension and must be included in assessments of language abilities. Children should be able to comprehend reversible active sentences by three years of age (de Villiers & de Villiers, 1973). Pointing out which picture shows the girl hitting the boy and which shows the boy hitting the girl is an example of a pair of reversible active sentences. Although this is extremely important in English, in many languages, including Spanish, word order is not a factor in meaning; and there is no equivalent test. The issue of what English Language Learners should be expected to know at these ages was discussed. Another feature of sentence structure that children should understand by four is the difference between the subject and object of a sentence in passive voice (e.g., "The boy was hit by the girl.") as opposed to active voice (e.g., "The girl hit the boy.") (de Villiers & de Villiers, 1973).

Semantics

Semantics refers to the knowledge of the meaning of words. The subgroup focused on not just the understanding of lists of words, but also the concept development that is associated with those specific meanings. For example, the understanding of quantifiers, such as "all" and "none," is established by four years of age, and is important not only in language, but also for mathematics skills. This means that semantics has the potential to play a role in the assessment of both language and some aspect of mathematical ability. Recent research from several groups, particularly at the University of Maryland, has discussed the importance of these semantic concepts, and how they relate to pragmatics and syntax (Crain & Thornton, 1998; Lidz & Musolino, 2002; Philip, 1995).

Pragmatics

Pragmatics is an area that is important in both language comprehension and expression but is often overlooked in standardized tests. Aquiles Iglesias has developed a "birthday task" that measures many of these characteristics in one sitting (See also the Integrative Assessment Methodology section.). Given the opportunity to design a new instrument, Iglesias' task was considered a good starting point from which to cover all aspects of language development at these ages (Pena, Iglesias, & Lidz, 2001).

By three years of age, children should be engaging in *appropriate speech acts*, such as asking questions, clarification, denying, describing, and naming. One example from Iglesias' test is the "gift-wrapping" task. The experimenter asks the child to do several things that are not possible, do not make sense, or need to be further explained. Children should be able to engage in speech acts, such as asking and answering appropriate questions with regard to what they do not understand. It was suggested that other issues of social and emotional development can be tested easily in the same context (e.g., creating situations of frustration, creating a delay in gratification).

Iglesias is confident that assessments such as this can be predictive of later outcomes and can be used cross-culturally. These tests are appropriate for ages three and four, with some additional abilities expected of four-year-olds. For example, four-year-olds should be able to convey the appropriate speech of others, such as explaining what Jim needs to do if he wants one of Mary's cookies. However, special consideration to possible cultural differences needs to be applied for this type of measure.

Narrative

Another important test of semantic/pragmatic ability is the *cohesion of a story*. The ability to use narrative is important because it is related to reading ability (Snow & Dickenson, 1990). This can be tested in an elicited narrative in which children are asked to tell a story. Within this story,

competencies such as *staying on task*, the use of *goals and mental state words*, and *reference specification* would be important to assess. For example, does the child *differentiate characters* so that the audience can follow the storyline? These abilities are important in both expressing stories and in understanding them. Some of these measures are also informative with respect to social development and would therefore provide an opportunity for integrating the assessment of multiple domains in one task (de Villiers, 1989).

Literacy

The subgroup supported the use of existing measures that test literacy knowledge and development, and agreed that these abilities are important for school readiness. For three-year-olds, these tasks include *print concepts* such as book handling, page turning, and differentiating pictures from words.

By age four, the measure should become much more difficult, including recognizing letters in the child's name (as opposed to knowing at least 10 letters), writing some letters, and some letter-sound pairing (e.g., "Which letter makes the 'kaa' sound?"). It was noted that these abilities are not equally important to all cultures. For example, in many areas of Europe, the actual names of letters are not taught. More often, just the sounds associated with letters are taught to young children. Lastly, children should be able to recognize some whole words, such as their own name, "Stop," and "Exit."

Assessment Instruments

The subgroup identified the following competencies as those that are most commonly assessed.

- Increase vocabulary
- Acquire English
- Concept of a word
- Print awareness
- Phonemic awareness
- 10 letters of the alphabet
- Relationship of letters and sounds in writing

A discussion of various assessments for these and other competencies followed.

Inappropriate Instruments

The flaws of several measures of the above constructs were discussed. For example, it was noted that the most recent edition of the *Peabody Picture Vocabulary Test (PPVT-4)* does not exist in Spanish. The *Bayley Scales of Infant Development (BSID-II)* is not available in Spanish either. The *Woodcock-Johnson – Sound Awareness* (subtest) is another commonly used measure; however, its validity is weaker for children as young as four years. Finally, the *Test of Early Readiness Ability (TERA-3)* was well liked except for being a poor measure of phonological awareness.

Recommendations

The language measures that were recommended by the subgroup, until better assessments can be created, included

1. *Preschool Language Scale-Third Edition (PLS-4*) (Zimmerman, Steiner, & Pond, 2002), Auditory subtest, and
2. *Expressive One-Word Picture Vocabulary Test (EOWPVT*) (Brownell, 2000).

The auditory subtest of the *Preschool Language Scale* was considered favorable because it is available in Spanish and is a short (10-12 minutes) and valid test. The subtest can be interpreted on its own and is a good measure of vocabulary and other competencies. In addition, for the

purposes of testing growth in English Language Learners, Susan Levine has found strong results showing growth in English for native Spanish speakers using the auditory *Preschool Language Scale*. The *Expressive One-Word Picture Vocabulary Test* was chosen because of its validity regardless of children's native language. Together these tests are estimated to take only 20 minutes to administer.

Aware that the following recommendations for literacy would take far too long to administer all together, the group offered several options, from which decision makers can pick.

1. *Test of Early Reading Ability (TERA-3)* (Reid, Hresko, & Hammel, 2001) for print concepts
2. *Developing Skills Checklist* (CTB-McGraw Hill, 1990), Auditory Processing Subtest for phonological awareness
3. *Dynamic Indicators of Basic Early Literacy Skills (DIBELS)* (Good, 2000) for letter knowledge
4. *Get Ready to Read* screener *(GRTR)* (Whitehurst & Lonigan) for word concept – only if teachers are chosen to be the administrators of testing

Although the *Get Ready to Read* screener only has 20 items, it has proven to be a valid measure of school readiness. Administration is recommended for two to three times a year, and growth across that time span has been found. It contains pointing to the word/picture item, which all members of the subgroup agreed was very important. It also includes many other important tasks, such as letter-sound pairing (e.g., point to the one that makes the "sssss" sound) and tests of alliteration and rhyming. On the other hand, *GRTR* is not a good measure of phonemic awareness. Another drawback is that the instrument is a screener from which individual items cannot be pulled out separately, especially not for the purposes of national reporting. Because of these drawbacks, this screener is only recommended if teachers are required to conduct child evaluations.

The *Developing Skills Checklist - Auditory Skills Subtest* (CTB-McGraw Hill, 1990), which has been used by Whitehurst, was recommended as the best test available for the assessment of phonological awareness. The subgroup members agreed that this test is not ideal, but it is short and has reliable psychometric properties. It also contains items for early listening skills which would be important to test.

The *Dynamic Indicators of Basic Early Literacy Skills (DIBELS*) (Good, 2000) is a dynamic indicator that focuses on a number of behaviors thought to represent critical prereading skills. It has demonstrated strong reliability and validity in terms of its capacity to chart growth. The subgroup recommends the letter naming items, in particular.

<u>New Measures Developed by Participants</u>

Jill deVilliers is part of a team headed by Professor Harry Seymour at the University of Massachusetts conducting NIH funded research aiming to develop norms for children speaking African American English. Most existing tests are biased in terms of phonology and morphology towards Mainstream American English, which results in many African American children inaccurately classified as impaired. The new measure they have developed, the DELV (The Psychological Corporation, 2003) is based on contexts which are more universal and free of dialect-specific linguistic forms. Such an instrument needs to be non-biased but also sensitive enough to identify disorders. deVilliers claims that her measure is in some ways actually more difficult than more widely used instruments, contrary to criticism that such tests would end up being too easy. The test seems to be successful in that language-impaired children are distinguished from normally developing children without showing a bias against African American English speakers.

The DELV includes subsections for Morphology and Phonology (adjusted to be neutral for dialect), Syntax, Semantics and Pragmatics, and includes many innovative tasks based on contemporary

research in language acquisition. However, as it exists now, it may not be the best suited instrument for the needs of most preschool assessment. It is a 45-minute test which was developed as a clinical instrument to screen children who have language disorders, not an educational assessment instrument. However, it must be pointed out that most existing language tests also were developed under that clinical, not educational, goal. Instruments to assess the adequacy of educational interventions, or to rate children with respect to each other in the normal range, is outside the purview of existing tests for preschoolers.

Mathematical Skills

The Mathematics/Spatial Skills subgroup noted that early mathematics concepts go beyond learning to count and acquiring a few number facts (Ginsburg, Klein, & Starkey, 1998). In this context, the group found the term "Numeracy," which is sometimes used as a counterpart to "literacy," to be problematic, because the term does not encompass the non-numerical aspects of preschoolers' mathematics, such as spatial and geometric concepts and skills. "Mathematical Literacy" was suggested as a better term. This perhaps implies an important analogue to a goal that has been commonly promoted for the preschool age group in the language domain. Just as we encourage not only phonics and comprehension but also a love of reading, so we need to foster not only mathematical concepts and skills, but also a love for mathematics.

Important Constructs to Assess

Although not providing a complete list, the subgroup identified the following constructs as among those important for understanding mathematics development.

1. Counting
2. Enumeration
3. Mental "number line"
4. Addition/subtraction
5. Commutativity of addition
6. Simple plane and solid shapes
7. Pattern
8. Spatial relations
9. Symbolization
10. Estimation
11. Graphs
12. Telling stories about numbers
13. Use of language to describe problem-solving.

In almost all cases, participants stressed the importance of assessing not only content, but also solution process, as well as metacognitive and representational aspects of mathematics ability (e.g., creating stories about mathematical problems and describing methods of calculation).

By way of example, the group analyzed the academic standards for kindergarteners set by the Chicago Public Schools (http://intranet.cps.k12.il.us/Standards/CAS/cas.html), which call for children to be able to do the following tasks.

1. Count, read, write, and order numbers to 100
2. Compare whole numbers up to 100 using words more than, same as, less than
3. Represent numbers using physical models

4. Recognize number words through ten
5. Represent number facts (sums) through 20
6. Recall number facts (sums) through 20
7. Add 2-digit numbers
8. Solve problems that involve addition
9. Demonstrate the use of symbols for addition (+) and equal to (=)

The subgroup evaluated this list and found it to be too concrete and failing to focus on mathematical ideas. It was criticized for not examining deeper knowledge or emphasizing a love of mathematics. Another specific criticism pertained to the arbitrary cut-offs. For example, number facts are better described in terms of learning specific and sometimes meaningful number "families" (e.g., the "+1" family or the doubles) than learning isolated number facts up to some limit, such as 20.

Shape and Pattern

Participants discussed the importance of understanding shape (Clements, 1999). Both shape and pattern have been found to be predictive of later child outcomes. Using their PTCS measure, Janellen Huttenlocher and Susan Levine have found pattern sequencing ability to be an important cognitive skill. Doug Clements has also found his measures of shape composition and pattern sequences to be predictive of later learning. He also stressed that these skills were heavily input sensitive. With regard to future assessments of these constructs, there was a suggestion to find ways to test shape knowledge independent of vocabulary (e.g., a task which asks: "Show me something that goes with these things."). In addition, it was suggested that examining symmetry would be important in future tests of shape.

Space and Mapping

Members of the subgroup felt strongly that the use of graphs and mapping should be included in preschool assessment. Susan Levine shared her data on 4-year-olds regarding map usage. In addition, David Uttal showed a demonstration of a symbolic mapping task that he has been developing. The task taps into the symbolic reasoning that emerges between 4 and 5 years of age. It is a small-scale, pen-and-paper test, but it has demonstrated developmental correspondence to the use of maps on a larger scale (e.g., in room searches). Uttal outlined a trajectory of what children should be able to do by first grade in terms of mapping and suggested a set of spatial reasoning tasks which he recommends for inclusion in a preschool assessment battery.

The Role of Teachers

The role of teachers in mathematics learning was discussed at some length. Susan Levine has conducted a study that analyzed teachers' math-related language (e.g., spatial language during story time). She found large variability among the teachers, which was not related to social class. The teacher's math-related language was correlated with children's growth on certain measures. Also, lower SES children performed less well compared to higher SES children no matter what the measures. These children are estimated to be more than a year behind their more advantaged peers when they begin kindergarten.

What skills do teachers find important to teach, and what abilities are teachers actually teaching when following a lesson plan? It was recognized that teachers have their own goals and views of math, which are different from those of researchers. Teachers' understanding of mathematics is sometimes limited and they may underestimate children's mathematical abilities. Data on teachers' interpretations of children's behaviors and judgments of good learning have been collected by Herbert Ginsburg's "Big Math for Little Kids" program (Ginsburg, Greenes, & Balfanz, 2003. He found that teachers sometimes have misguided judgments about effective methods of teaching.

Assessment Instruments

The Role of Naturalistic Observation

Some participants have shown that researchers can gain insight into children's mathematical knowledge by observing their free play. In particular, free play was considered a useful situation in which to study children's work with shape, space, and symmetry (as revealed for example in block building). Joan Stiles' research in this area with brain-damaged and normal children was mentioned. Herbert Ginsburg has conducted studies employing naturalistic observation of potential SES differences in preschool children's free play (Ginsburg, Lin, Ness, & Seo, in press; Seo & Ginsburg, 2003). The frequencies of behaviors such as enumeration, pattern and shape exploration, and magnitude were examined. Children exhibited these mathematics behaviors with relatively high frequency compared to categorization and spatial behaviors. Furthermore, no SES differences were found in these aspects of mathematical behavior.

The Role of Clinical Interview

To gain insight into children's thinking, it is often necessary to go beyond direct observation to ask probing questions (Bowman, Donovan, & Burns, 2001; Ginsburg, 1997). In another study, Ginsburg and colleagues conducted clinical interviews of children's number concepts (Pappas, Ginsburg, & Jiang, in press). SES differences were not found in the use of strategies, but did exist for metacognitive abilities, such that the upper-income children were able to explain their answers more fully than were the other children.

Nonverbal Assessment

Because language becomes a limiting factor in the assessment of all developmental domains, the subgroup stressed the potential importance of non-verbal assessments.

Recommendations

For assessment of mathematics and spatial skills, this subgroup recommended

1. Subtests of the *Primary Test of Cognitive Skills (PTCS)* (Huttenlocher & Levine, 1990) and
2. *Test of Early Mathematics Ability-Third Ed. (TEMA-3)* (Ginsburg & Baroody, 2003).

The group also recommended not using the *Woodcock-Johnson* for the assessment of mathematical skills in preschoolers because it has not been validated for early childhood, nor is it based on current research of the development of mathematical thinking.

New Measures Developed by Participants

Test of Early Mathematics Ability (TEMA)

TEMA-3 (Ginsburg & Baroody, 2003) tests mathematics ability in children three to eight years old. The items of the instrument were chosen based on existing research and national norms. Almost every item is linked to an empirical research study. The items are sequenced in order of increasing difficulty. A separate Probes guide (Ginsburg, 2003) provides a series of follow-up questions to be used after the standard testing to examine children's methods of solution and their "zone of proximal development" with respect to key items failed during standard administration. For each item, the Probes session begins with re-worded questions designed to determine if the child did not understand the original question. A strategy question then follows to identify the child's method of solution (e.g., tell me what you are thinking about this problem?). Next, a justification question is asked (e.g., can you prove to me that 2 and 2 is 5??). Finally, the examiner gives a hint (e.g., how about using your fingers to count?) to determine whether the child can solve the problem with some adult assistance. Conducting this sequence of questioning may reveal the source of the child's difficulty, which could directly inform a teacher's instruction or a clinician's intervention. With NSF funding, Ginsburg is currently working on software that can guide teachers' mathematics assessment and organize their observations.

Building Blocks Mathematics Assessment
Douglas Clements' measure has been developed and field tested for various aspects of mathematics along research-based learning trajectories. The assessment uses an individual interview format, with explicit protocol and scoring procedures. It has two tests. The number component includes items measuring verbal counting (forward, backward, up from a given number, before/after/between; includes identifying mistakes in counting), object counting (counting groups in array and scattered arrangements, producing groups; includes identifying mistakes), subitizing, number comparison (nonverbal and verbal), number sequencing, connection of numerals to quantities, number composition and decomposition, adding and subtracting (including concrete situations, story problems, and mental arithmetic), and place value. Children proceed along research-based trajectories until they have made three consecutive errors. The final items measure skills typically achieved at eight years of age. The maximum score is 97; for this sample, children reached items associated with 6.5 years of age; therefore, the practical maximum was 78 (reliability, r = .89). The geometry test measures shape identification, shape composition and decomposition, congruence, construction of shapes, and turns (one item). It also included items on geometric measurement and patterning. As with the number component, difficulties for some items on the geometry assessment were designed to measure abilities at 8 years of age. Children complete all 17 items (seven of which have two parts), for a maximum score of 24 (reliability, r = .71, lower due to the disparate nature of the concepts assessed).

Primary Test of Cognitive Skills (PTCS)
Janellen Huttenlocher discussed the research that she and Susan Levine have conducted which involved testing large groups of preschool children using the PTCS to determine the children who were likely to experience difficulty in school later in life. The PTCS assessed nonverbal math, as well as memory, spatial ability, and vocabulary. Examining the correlates between each of the individual tasks showed a high correlation between syntax and vocabulary and a high correlation between comprehension of syntax and production of syntax. They also developed a comprehension scale. The sample included a large array of children from all backgrounds, which produced vast variability in syntactic skills. Children's ability for non-verbal mathematics was an excellent predictor of later special education needs. Huttenlocher and Levine are now following a large group of children to study these two variables. Huttenlocher presented a demonstration featuring some of the measure's comprehension questions. She mentioned that sometimes only a few items are needed to get a large amount of information about a child. She also mentioned the possibility of group testing, as well as its disadvantages.

Measure of Verbal Math
It is possible for a child to understand number concepts, but not number language. Susan Levine described her measure of verbal math, a task which involves guessing the number of items under a box. She has found social group differences on verbal mathematics but not non-verbal mathematics. However, teachers might construe children as knowing less because of verbal limitations. This research concurred with her earlier work with Huttenlocher with respect to the predictive power of nonverbal mathematics for later special education needs. In another study, Levine found that the amount of teacher input is highly correlated with children's verbal math, but not their non-verbal mathematic ability. Therefore, nonverbal mathematics needs to be assessed in order to identify the children who are conceptually behind. One the other hand, verbal mathematics needs to be examined in order to identify children who may not be in optimal environments. Levine stressed the value of being able to solve problems verbally. In later years, children from lower social classes do progressively worse on tasks that rely more heavily on verbal skills (such as problems with large objects sets).

Social-Emotional Skills

The subgroup generated an extensive list of social-emotional "competencies" with corresponding subcomponents. The term "competencies" was adopted instead of the traditional term "milestones." Competencies should be thought of in terms of inter-related skills, rather than the lockstep milestones theory of the past. Research shows that many social competencies emerge together during the preschool years, from three to five (Campbell, 2002; Eisenberg & Fabes, 1998; Saarni, Mumme, & Campos, 1998). The emerging social competencies include behaviors that reflect the child's individual functioning, including early interest in exploring and mastering the environment, and regulatory skills such as self control and emotion regulation. Social development at this age also includes a range of peer-related and social-cognitive skills that are reflected in the child's ability to relate to adults and other children, such as social awareness, perspective taking, prosocial behavior, and the ability to engage in sociodramatic play. In addition, it was noted that any consideration of the social and emotional competence needs to also consider the assessment of early behavior problems (Campbell, 2002).

Important Constructs to Assess

Regulatory Skills

Basic regulatory skills include attentional control, listening skills, following directions, learning rules, and planning. Emotion regulation consists of controlling or managing negative affect and displaying appropriate expression of positive affect. Self control refers to the ability to resist temptation and distraction, tolerate frustration, and comply with requests. Exploration/mastery is the trait which represents the child's level of interest, curiosity, initiative, persistence, and motivation to explore. (See Calkins, 1994, Campbell, 2002, Eisenberg & Fabes, 1998; Rothbart & Bates, 1998; Thompson, 1994).

Prosocial Behavior

With regard to young children interacting with others, social awareness such as self-other differentiation and a sense of self come into play. Perspective taking is also an essential skill that emerges during the preschool years. Other important prosocial behaviors include initiating interactions with peers, cooperating with others, sharing, turn-taking, working toward group goals, showing concern for others, and resolving conflicts without resorting to aggression. Similar skills are needed for the mastery of sociodramatic play (e.g., shared goals and turn taking), but sociodramatic play also requires representational ability, role assignment, and shared scripts. (See Campbell, 2002; Dunn, 1988; Eisenberg & Fabes, 1998; Rubin, Bukowski, & Parker, 1998; Saarni et al., 1998).

Prosocial skills with adults are also necessary for preschoolers' optimal learning. For example, children need to know how to use adults appropriately to support their social interactions and to support their learning. Children who feel competent and comfortable are more likely to seek the help of an adult to resolve peer disputes and to scaffold their problem solving. By the same token, children need to recognize the situations in which adults should not be sought to intervene. Robert Pianta's (1999) research has demonstrated the importance of teacher-child relationships for child outcomes, a relationship which undoubtedly depends on a child's ability to use adults appropriately.

Behavior Problems

Social competence is defined in terms of the emergence of self control, dramatic play, self-awareness, and social-emotional understanding. Problems in these areas may be signs of early-appearing behavior problems, for example, when children have difficulty regulating their behavior, playing cooperatively with others, or following directions. Although there are wide individual differences in children's social competence in the preschool period, some children at the extremes

of over- or under-control, or at the extremes of social withdrawal or social explosiveness may be showing early signs of behavior problems (Campbell, 2002).

Problem behaviors may be described in three different patterns:

1. Under-control problems (e.g., aggressive children)
2. Over-control problems (e.g., socially withdrawn children)
3. Developmental issues (e.g., toilet training).

Other behavioral concerns (such as difficulties with toilet training, separation anxiety, food fads) may merely reflect immaturities and uneven development, rather than potential indicators of problems. Furthermore, there are differences in the rates of behavior problems which are related to child gender. These differences have serious implications for later development. Given the strong link between behavior problems and learning problems, it becomes important to assess social-emotional functioning in order to ensure school success.

Determining what behavior is problematic could lie partly in the eyes of the beholder. For example, a teacher with a high tolerance for activity in the classroom may interpret "externalizing" behaviors differently from a teacher with a low tolerance for activity. Low tolerance may be one of the factors in the tendency to overpathologize children. There was mention of a recent report showing an increase in ADHD medication for children as young as three years, which may be one example of this. On the other hand, in light of evidence indicating that behavior problems often are difficult to treat when they are not caught early, and the evidence that interventions implemented early in a child's life lead to the most gains, a strong case can be made for early interventions when appropriate (see Campbell, 2002 for a review of these issues).

Temperament

Temperament is an important and interesting child characteristic because it is considered the biological basis of individual differences (Rothbart & Bates, 1998). Temperament includes such dimensions as physiological reactivity and regulation, as well as deploying attention and effortful control. Nathan Fox's longitudinal research, which has followed trajectories as they emerge from reactive biases as infants to thinking about more complex issues in later childhood, shows a profound influence of temperament on cognitive development (see Fox, Henderson, & Marshall, 2001 for a review). He has found that behaviorally inhibited and fearful children are inflexible and have more difficulty learning.

A Whole Child Approach

The tendency for assessment to carve the child up into different developmental domains is problematic and has resulted in misconceptions about development. The subgroup discussed the need to see the child as an integrated whole in order to understand the influences of cognition on social and emotional behavior and of social-emotional behavior on cognition and learning. In essence, there is little possibility of cognitive advances without the ability to direct/focus attention and regulate emotions.

Assessment Instruments

Measurement Issues

The whole group discussed social-emotional assessment further and noted the possible non-linear aspects of some social and emotional behaviors, as well as the importance of the context of assessments. Depending upon the behavior, frequency of occurrence may not be the only aspect of interest for assessment. For example, the intensity of a behavior may matter, even if it only happens once. That is, children may not aggress often, but when they do, it is important. In this case, pass/fail type items may be more appropriate as a valid measure. On the other hand, in

assessing social competence/self-regulation it may not be enough to know only whether the child has the skill or not, as is sometimes the case in the language domain. It is more valuable to determine whether the child can utilize the appropriate social regulation skills in demanding situations requiring restraint or awareness of the needs of others.

Furthermore, much of social-emotional behavior is context specific, more so than language, spatial or mathematic abilities. Context-specific behaviors are difficult to measure with a single standardized measure. In addition, it was recognized that the direct assessment of social skills may be confounded by the child's language ability and understanding of the instructions. Given these difficulties, the merits of parent/teacher reports of children's social-emotional development were discussed.

Teacher/Parent Report Measure

Participants suggested the use of both parent and teacher report in order to address the issue of context. Adult reports are necessary especially for data on extreme behaviors such as behavior problems. It was recognized by all, however, that it is more difficult for teachers than for parents to accurately report and be knowledgeable with respect to ranges of normal behavior. Nonetheless, parent/teacher report is necessary not only for identifying behavior problems, but also for the assessment of typical social development.

Another issue in teacher report of children's social behavior is that significant biases may be reflected. With regard to accountability, would teachers accurately report on children's behavior as part of a process in which their own programs are being evaluated? Thus, different measures will be needed depending on whether the goal is to evaluate programs or to predict a child's future success. In order to eliminate teacher bias, one idea was to ask teachers about how they think the children will be doing in the future rather than how children behave now with them. Another idea was for developing a teacher/parent report measure with a "lie scale" similar to that of the MMPI. In any case, the use of teacher reports of children's competencies will require testing theory about teacher bias related to changing age appropriate behavior.

Using Measures from Social-Cognitive Neuroscience

The participants felt strongly about prioritizing self-regulation, in particular, in preschool assessment because of the role it plays throughout the lifespan, and therefore its potential to be a public health issue. There was some discussion about using the broader term of "executive function" rather than self-regulation. The literature on executive function overlaps with the temperament literature such that the terminology has become confusing (e.g., terms such as response inhibition, adaptive behavioral control, and effortful control). From the cognitive neuroscience literature, there is evidence that executive functioning has different components that are broader than just self-regulation.

A discussion of neuropsychological testing noted that tasks that were first developed in the adult cognition literature are now being adapted for preschoolers (e.g., sun-moon version of the Stroop Test) (Perez-Edgar & Fox, 2003). Some believed that most of these types of tasks are not yet well developed for children (e.g., non-verbal form of Stroop Test). These tasks do have potential because they tap into some underlying abilities, but they are de-contextualized and therefore should be interpreted with caution. There would still be a need to examine larger-scale behaviors and behavior within and across contexts. Once again, the need to rely on adult reports for certain types of preschool behavior was raised.

In addition, researchers have been adding affective components to what otherwise have been more "cognitive" tasks. For example, Nathan Fox has added a win-lose point component to the Posner task. Although it still remains more of a cognitive task despite this change, it seems to be a promising merger of affective and cognitive measures. Another example was the "Flanker task" in

which subjects tend to slow down reaction time after making errors. Nathan Fox has measured ERP during such tasks. In 7-year-olds, he has found a relationship between performance on this task and behavioral regulation in social situations.

The question remained as to whether neuropsychological results are predictive of real-world behaviors. Some evidence from adult neuropsychology in which impaired persons perform well suggests that neurological measurement may not be ecologically valid with regard to developmental preschool assessment. It was suggested that computer-based methods are assessing not social behavior, but underlying constructs that are related to social behavior. Furthermore, it is difficult to get individual differences for young children because individual gradients do not exist. Thus, while the group felt that measures from social-cognitive neuroscience hold promise, more research would be necessary before they could be recommended.

New Measurement Approaches

There was a discussion of how assessment of social development might be combined with existing assessment of cognitive functioning. One suggestion was to embed the measurement of some aspects of social behavior in other assessment procedures. For example, behavioral observations could be scored during psychological testing. However, others provided examples of well-adjusted, socially competent children who may only experience problems during testing situations. This suggests that the testing context may be too limited for determining a child's overall social-emotional functioning and competency.

Recommendations

Social emotional functioning is relevant for school readiness, measurable, and likely to be affected by curricula or programming. The group tried to pick out the "learning-related social skills," but found this a difficult task. Constructs such as response inhibition, working memory, attention switching, delay of gratification, and error monitoring were offered; however, there was some concern regarding whether these were the most valid representation of social-emotional behavior.

The constructs that the subgroup recommends as important to measure in preschool settings are self-regulatory skills, prosocial behavior with peers (sharing, cooperation, turn-taking), prosocial behavior with adults, and emotional regulation. It was suggested that a simple checklist such as the *Preschool Behavior Checklist (PBR*) (McGuire & Richman, 1993) be used as an observational component of assessment. Also, a parent report instrument, serving as a "customer satisfaction measure" would also be important for preschool program evaluations.

A multi-faceted approach was recommended in which the following three things could converge:

1. Measurement of social skills embedded in other assessments,
2. Behavior rating during assessments of other developmental domains, and
3. Teacher report for a broad spectrum report of competencies.

Although the subgroup did not come to a consensus decision with regard to recommending specific existing measures, those that the subgroup thought showed some promise include

1. the *Social Competence/Behavior Evaluation Short Form (SCBE-30*) (LaFreniere & Dumas, 1995),
2. the social measures used in the Early Childhood Longitudinal Study (ECLS – West, et al.),
3. a shorter version of the *Q-Sort* for attachment (Waters & Deane, 1985), and
4. the *Devereux Early Childhood Assessment (DECA*) (LeBuffe & Naglieri, 1999).

PART THREE: FUTURE DIRECTIONS IN PRESCHOOL ASSESSMENT

The Need for New Research

The need for more research on assessment is well established. In the early 1990's, when the National Education Goals Panel proposed that "all children will start school ready to learn by 2000," the absence of means to assess dimensions of school readiness, such as language usage, approaches to learning, and cognition, limited reports of child well-being to indicators such as low birth weight, immunizations, and attendance at high quality preschool. In a meetings sponsored by federal funding agencies, researcher Jerry West outlined the challenges he and his colleagues experienced in finding appropriate measures for the Early Childhood Longitudinal Study, on which validation work is still needed. He emphasized the urgent need for new measures with sufficient breadth and depth of coverage while remaining feasible in terms of time, cost, training and implementation on a large-scale. The National Children's Study now faces the same difficulties with respect to measurement for its large sample of infants.

Forum participants felt strongly about the need for funders to seriously consider investing in instrument development. Current and future studies are caught in a continued cycle: They need appropriate measures but do not have the time or money to develop them. An instrument development initiative would help put an end to this problem. If we strive for assessments that measure constructs deemed crucial by scholars in each domain, that are ecologically valid, integrative and unbiased, accumulation of converging evidence across studies would be possible. Data of such integrity would be powerful enough to inform issues of national significance, such as scientifically based curricula.

Overcoming Test Biases

Existing assessments suffer from serious questions regarding their validity for special populations. The issue of culture- and language-fairness is of paramount importance in addressing the testing of preschoolers. For example, among the children in Head Start, 139 languages are spoken. Many of the children in Head Start are speakers of Spanish or from homes where Spanish is spoken. Another large group comes from homes and neighborhoods where African-American English is the language variant used. The existing tests do not take these language variants into account, nor do they attend to the cultural and contextual differences in these children's experience. Most of the current tests evaluate mastery of mainstream English, and deal with cultural variation by meeting the criterion of "inclusion." The latter demands that the standardization sample must match the Census data. However, minority representation in the standardization sample does not address the possibility that minority children may not perform as well as majority children because of test bias, for example, whether the questions about knowledge are normed with respect to the contexts of minority children's lives (Adler & Birdsong, 1983; Kahmi, Pollock & Harris, 1996; Seymour,

Bland-Stewart, & Green, 1998; Seymour, Roeper & de Villiers, 2003; Stockman, 1996; 2000; Washington & Craig, 1990).

In the areas of IQ and language tests, it is well-established that these biases exist. The State of California has a law derived from the case of Larry P. v. Riles (1979), forbidding the use of standardized intelligence tests that are not normed on African American children to determine the eligibility of African American children for placement in an EMR setting "or its substantial equivalent." According to current California Department of Education guidelines, this ruling applies to all special education placements. In addition, it applies not only to IQ tests but to any tests (including standardized speech and language tests) that are validated against an IQ test. This would block many of the proposed tests from being used in California Head Start, which is 10% of the total (Affeldt, 2000; California Department of education, 1989; California Speech-Language-Hearing Association, 1994; Wyatt, 2003 Task Force report).

It can be argued that since these children are in classrooms designed to encourage the use of mainstream English and mainstream American values, the tests (although biased) can assess growth towards these societal goals. But when test results are aggregated to evaluate schools' effectiveness, these aggregate scores are potentially misleading because of the mismatch of school and home cultures and language. Furthermore, when tests mainly assess how children perform to a mainstream standard they draw our attention to deficiencies with respect to that standard, and draw our attention away from equally important information about their proficiencies. These may be equal to the mainstream population in language, problem solving, comprehension skills and other cognitive abilities. A vital need for addressing this issue is more research on instruments designed for the population in question, sensitive to their differences, and appropriate in their expectations.

The Need for Ecological Validity

Most assessments lack ecological validity, rendering them ineffective for informing teaching. They therefore have limited benefit for improving children's learning. Assessments that are process oriented would help teachers focus on children's learning and the environment in which instruction is taking place, as well as prevent them from "teaching to the test." Moreover, ecological validity would help obtain the goal of using test results to develop individual education plans for children prior to kindergarten.

Constructs Most Relevant for School Readiness

Existing instruments are not tapping the full range of constructs that child development researchers believe are scientifically most relevant for early school readiness and long-term scholastic success. Socioemotional development is in special need of instrument development, but even in well-defined areas such as language/literacy and mathematical development, the tests often assess outcomes without enough attention to the underlying processes that support these outcomes. Appropriate measures need to be developed to assess these constructs.

Researchers at the Temple Forum reached consensus about important language and literacy constructs to measure, not all of which seem to be on current lists (such as the federally mandated outcomes for the Head Start program): Vocabulary diversity, Narrative use and understanding, Rhyming, Syntactic complexity, Alliteration, and Quantification (e.g., "each and every, one and only"). Beyond these, scientific evidence shows that to be prepared for school, children also need to be developing numeracy skills, strong problem-solving and other cognitive skills, and social competency. Thus, the group recommended that the design of the assessment system be

consistent with Head Start's school readiness goals and cover these essential domains from its inception.

Integrative Assessment Methodologies

There are very few integrative assessment procedures that allow for a comprehensive understanding of the nature of children's learning, and that adequately evaluate how competencies in different developmental domains interact to produce optimal functioning and school readiness. For example, a child's progress toward reading and mathematic proficiency in preschool depends on the ability to regulate attention and to use language flexibly in the service of multiple goals. Similarly, progress in social skills and social competence requires the regulation of emotion and the development of a sense of self as an efficacious and active learner. Ultimately, the integration of social and academic competence in the regulation of attention and emotion can serve as the basis for ongoing achievement in school. Thus, an integrated assessment of cognitive and social development is important if there is to be progress in the field of preschool assessment.

Methodologically, one way of addressing integrated assessment of developmental domains is through dynamic assessments conducted in comfortable, familiar settings that are of interest to the child. Such settings enable preschoolers to better demonstrate their competencies. In doing so, strengths in cultural differences can be revealed, and the cultural biases which result from more rigid testing can be eliminated. Moreover, dynamic assessments are more ecologically valid because they incorporate aspects of the classroom setting to a much larger extent than conventional testing, rendering the results far more useful to teachers for their curriculum planning.

Dynamic and integrative assessment methods address another weakness found in most conventional testing: Heavy reliance on children's language abilities. Although measures that require the examiner to elicit responses from the child can be effective for assessing some concepts about language and mathematics, a profoundly different view is provided when the child is the initiator of activities. Language that is used functionally in everyday contexts is likely to provide richer information than elicited language which may not be representative of the child's functioning.

The ideas generated for specific dynamic and integrated assessment tasks were considered some of the most fruitful. Many ideas were voiced, and several of the participants who have developed or are developing these types of tasks shared their rewarding experiences using such assessment methods. The context of a birthday party has been common for many integrated tasks and is considered a culturally universal scenario. For studies assessing language delays in Latino children, Aquiles Iglesias has developed a successful birthday party task in which all questions are contextualized by being embedded within a story. In one feature of the birthday party task, the child is told that a "mushky" is in the box, and s/he is asked to describe how to wrap it as a birthday gift (Peña, et al., 2001).

The group expanded upon the concept of a birthday party task with examples of how assessments of language skills, social competence, delay of gratification, and numeracy (e.g., counting the number of gifts) could be integrated. For example, various language tasks could be embedded in the birthday party such as asking questions to the children with various levels of syntax, including a word game, and observing for word diversity. Nathan Fox has used a birthday party task in which four children participate, providing a context for the observation of social competence with peers. Ginsburg used a birthday party task in a cross-cultural comparison of mathematical abilities (Ginsburg, Choi, Lopez, Netley, & Chi, 1997).

Sampling Strategies for Accountability

Although gauging every child's progress is needed to guide teaching, when assessment is for the purpose of accountability, it is important to establish an effective sampling strategy. A random sampling process, in which a subset of children in classrooms is examined, would produce more in-depth information and more valid knowledge in the amount of time that will be available for testing children due to the limited duration that is appropriate for children of this age.

Also, matrix sampling is an efficient and effective way to collect valid data regarding accountability. Matrix sampling involves giving parts of tests to all children, rather than giving every child all of the test items. This sampling strategy would lower the risk of teachers teaching to the test since it would be impossible to know on which particular items any one child would be tested.

Conclusions and Action Steps

As a collaboration of developmental scientists who are invested in informing public policy regarding early education, the forum participants felt strongly about communicating the consensus reached at the Temple Forum. Disseminating a report of the forum proceedings was agreed to be an important way to share the group's recommendations. Immediately after the forum's conclusion, however, because the federal Head Start program was beginning to develop a standardized assessment battery at the time the forum was held, participants drafted a letter to Dr. Craig Ramey, one of the Chairs of the Head Start Technical Advisory Group. The letter urged the advisory group to consider the recommendations for existing measures and for sampling techniques resulting from the discussions at the Temple Forum.

The forum participants also recognized the potential for assisting the National Children's Study, which was in the beginning stages of selecting the best available assessments for the various developmental domains, by making similar recommendations based on evidence. Earlier drafts of this report were shared with key planners of the study. In addition, a recommendation was made for providing an opportunity for new instrument development as part of the funding allotted for this large national study.

As a more long-term goal, the group aims to inform funding priorities in evaluation research and the assessment of early child development and learning. For example, a letter to Duane Alexander of NICHD was drafted by the group to express the need for long-range planning to validate more culturally neutral assessments, some of which are already in development. The letter also explained why it is imperative to develop new assessments that are more ecologically valid and that help teachers identify those areas in which students need to grow and improve.

The Temple Forum was an initial step in addressing the problematic field of preschool assessment with only a small group of the scholars who conduct research on this important topic. The forum was successful in bringing the issues to the table and providing initial recommendations. More evidence needs to be examined with regard to all of the possible measures that may be appropriate to use with preschoolers, and more information is needed about instruments currently undergoing development. The resources collected through this collaboration will serve to refine the recommendations for the best existing assessments and needs for new assessments. The participants recognize the importance of developing effective ways of communicating this research to policymakers. CIRCL hopes to aid this process by building partnerships and developing dissemination strategies.

REFERENCES

Adler, S., & Birdsong, S. (1983). Reliability and validity of standardized testing tools used with poor children. *Topics in Language Disorders, 3(3)*, 76–87.

Affeldt, J. (2000). Legal issues in testing African Americans. In W. A. Thomas (Ed.), *Larry P. Revisited: IQ testing of African Americans*. (pp. 23-27). San Francisco, CA: California Publishing Co.

Anglin, J. M. (1970). *The growth of word meaning (Research Monograph No. 63)*. Cambridge, MA: The MIT Press.

Bartsch, K. & Wellman, H.M. (1995). *Children talk about the mind*. New York: Oxford University Press.

Bowman, B. T., Donovan, M. S., & Burns, M. S. (Eds.). (2001). *Eager to learn: Educating our preschoolers*. Washington, DC: National Academy Press.

Brown, R. (1973). *A first language: The early years*. Cambridge, MA: Harvard University Press.

Brownell, R. (2000*). Expressive One-Word Picture Vocabulary Test Manual*. Novato, CA: Academic Therapy Publications.

Bruner, J. (1986). *Actual minds, possible worlds*. Cambridge, MA: Harvard University Press.

California Department of Education (1989). *Larry P. task force report: Policy and alternative assessment guideline recommendations*. Sacramento, CA: Author, Special education division.

California Speech-Language-Hearing Association (1994). *CSHA Position paper: Assessment of the African American child.* CSHA, 21(3), 1-17.

Calkins, S. D. (1994). Origins and outcomes of individual differences in emotion regulation. In N. Fox (Ed.). *The development of emotion regulation. Monographs of the Society for Research in Child Development, 59*, 53-72.

Campbell, S. B. (2002). *Behavior problems in preschool children: Clinical and developmental issues*. New York: Guilford Press.

Carey, S. (1978). The child as word learner. In M. Halle, J. Bresnan & G. Miller (Eds.), *Linguistic theory and psychological reality*. Cambridge, MA: MIT Press.

Carey, S. (1982). *Semantic development: The state of the art*. In E. Wanner & L. Gleitman (Eds.), *Language acquisition: The state of the art*. New York: Cambridge University Press.

Carey, S. & Bartlett, E. (1978). Acquiring a single new word. *Papers and Reports on Child Language Development, 15*, 17-29.

Clark, E. (1993). *The lexicon in acquisition*. Cambridge University Press.

Clements, D. H. (1999). Geometric and spatial thinking in young children. In J. V. Copley (Ed.), *Mathematics in the early years* (pp. 66-79). Reston, VA: National Council of Teachers of Mathematics.

Clements, D. H., & Sarama, J. (2003). *Interim Report of Building Blocks—Foundations for Mathematical Thinking, Pre-Kindergarten to Grade 2: Research-based Materials Development (NSF Grant No. ESI-9730804)*. Buffalo, NY: State University of New York at Buffalo.

Crain, S., & Thornton, R. (1998). *Investigations in universal grammar: A guide to experiments on the acquisition of syntax and semantics.* Cambridge, MA: The MIT Press.

de Villiers, J.G. (1996) Defining the open and closed program for acquisition: The case of wh-questions. In M. Rice (Ed.) *Towards a genetics of language*. Hillsdale, N.J.: Lawrence Erlbaum.

de Villiers, J. G., & de Villiers, P. A. (1973). Development of the use of word order in comprehension. *Journal of Psycholinguistic Research, 2(4)*, 331–341.

de Villiers, J.G. & de Villiers, P.A. (2000) Linguistic determinism and false belief. In P. Mitchell & K. Riggs (Eds.) *Children's Reasoning and the Mind*. Hove, U.K.: Psychology Press.

Dunn, J. (1988). *The beginnings of social understanding*. Cambridge, MA: Harvard University Press.

Eisenberg, N. & Fabes, R. A. (1998). Prosocial development. In W. Damon (Series Ed.) and N. Eisenberg (Vol. Ed.), *Handbook of child psychology, Vol. 3. Social, emotional, and personality development*. (5th ed., pp. 701-778). New York: Wiley.

Fox, N. A., Henderson, H. A., & Marshall, P. J. (2001). The biology of temperament: An integrative approach. In C. A. Nelson & M. Luciana (Eds.), *Handbook of developmental cognitive neuroscience* (pp. 631-646). Cambridge, MA : MIT Press.

Ginsburg, H. P. (1997). *Entering the child's mind: The clinical interview in psychological research and practice*. New York: Cambridge University Press.

Ginsburg, H. P. (2003). *Assessment probes and instructional activities for the Test of Early Mathematics Ability-3*. Austin, TX: Pro Ed.

Ginsburg, H. P., & Baroody, A. J. (2003). *The test of early mathematics ability: Third edition*. Austin, TX: Pro Ed.

Ginsburg, H. P., Choi, Y. E., Lopez, L. S., Netley, R., & Chi, C.-Y. (1997). Happy birthday to you: Early mathematical thinking of Asian, South American, and U.S. children. In T. Nunes & P. Bryant (Eds.), *Learning and teaching mathematics: An international perspective*. (pp. 163-207). Hove (UK): Psychology Press.

Ginsburg, H. P., Greenes, C., & Balfanz, R. (2003). *Big math for little kids*. Parsippany, NJ: Dale Seymour Publications.

Ginsburg, H. P., Klein, A., & Starkey, P. (1998). The Development Of Children's Mathematical Thinking: Connecting Research With Practice. In I. Sigel & A. Renninger (Eds.), *Handbook of Child Psychology: 5th Ed., Vol. 4. Child Psychology and Practice*. (5th Ed. ed., Vol. Vol. 4, pp. 401- 476). New York: John Wiley & Sons.

Ginsburg, H. P., Lin, C.-L., Ness, D., & Seo, K.-H. (in press). Young American and Chinese children's everyday mathematical activity. *Mathematical Thinking and Learning*.

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition, 1(1)*, 3–55.

Huttenlocher, J., & Levine, S.C. (1990). *Primary Test of Cognitive Skills: Examiner's Manual*. Monterey, CA: CTB/McGraw Hill.

Kamhi, A. G., Pollock, K. E., & Harris, J. L. (Eds.). (1996). *Communication development and disorders in African American children: Research, assessment, and intervention*. Baltimore: Paul H. Brookes.

LaFreniere, P. J., & Dumas, J. E. (1995). *The Social Competence and Behavior Evaluation—Preschool Edition*. Los Angeles, CA: Western Psychological Services.

Larry P. v. Riles 495 F. Supp. 926 (ND Cal.1979).

LeBuffe, P. A., & Naglieri, J. A. (1999). *Devereux Early Childhood Assessment Program*. Lewisville, NC: Kaplan Press.

Lidz, J., & Musolino, J. (2002). Children's command of quantification. *Cognition 84*, 113-154.

Pappas, S., Ginsburg, H. P., & Jiang, M. (In press). SES differences in young children's metacognition in the context of mathematical problem solving. *Cognitive Development*.

Pena, E., Iglesias, A., & Lidz, C. S. (2001). Reducing test bias through dynamic assessment of children's word learning ability. *American Journal of Speech-Language Pathology, 10(2)*, 138–154.

Perez-Edgar, K., Fox, N. A. (2003). Individual differences in children's performance during an emotional Stroop task: A behavioral and electrophysiological study. *Brain and Cognition, 52*, 33-51.

Philip, W. (1995). *Event quantification in the acquisition of universal quantification*. Unpublished doctoral dissertation, University of Massachusetts, Amherst.

Pianta, R. C. (1999). *Enhancing relationships between children and teachers*. Washington, DC: American Psychological Association.

Rice, M., Buhr, J. C., & Nemeth, M. (1990). Fast mapping word-learning abilities of language delayed preschoolers. *Journal of Speech and Hearing Disorders, 55*, 33–42.

Roeper, T., & de Villiers, J. (1993). The emergence of bound variable structures. In E. Reuland & W. Abraham (Eds.), *Knowledge and language: Vol. 1. From Orwell's problem to Plato's problem* (pp. 105–139). Boston: Kluwer Academic Publishers.

Rothbart, M. K. & Bates, J. E. (1998). Temperament. In W. Damon (Series Ed.) and N. Eisenberg (Vol. Ed.), *Handbook of child psychology, Vol. 3. Social, emotional, and personality development*. (5th ed., pp. 105-176). New York: Wiley.

Rubin, K. H., Bukowski, W., & Parker, J. G. (1998). Peer interactions, relationships, and groups. In W. Damon (Series Ed.) and N. Eisenberg (Vol. Ed.), *Handbook of child psychology, Vol. 3. Social, emotional, and personality development*. (5th ed., pp. 619-700). New York: Wiley.

Saarni, C., Mumme, D. L., & Campos, J. J. (1998). Emotional development: Action, communication, and understanding. . In W. Damon (Series Ed.) and N. Eisenberg (Vol. Ed.), *Handbook of child psychology, Vol. 3. Social, emotional, and personality development*. (5th ed., pp. 237-310). New York: Wiley.

Sarama, J., & Clements, D. H. (2002). Building Blocks for young children's mathematical development. *Journal of Educational Computing Research, 27(1&2)*, 93-110.

Scarborough, H. (1990). Index of productive syntax. *Applied Psycholinguistics, 11*, 1-22.

Seo, K.-H., & Ginsburg, H. P. (2003). What is developmentally appropriate early childhood mathematics education? Lessons from new research. In D. H. Clements & J. Sarama & A.-M. DiBiase (Eds.), *Engaging young children in mathematics: Findings of the 2000 national conference on standards for preschool and Kindergarten mathematics education* (pp. 91-104). Hillsdale, NJ: Erlbaum.

Seymour, H. N., Bland-Stewart, L., & Green, L. J. (1998). Difference versus deficit in child African American English. *Language, Speech, and Hearing Services in Schools, 29(2)*, 96–108.

Seymour, H., Roeper, T., & de Villiers J. (2003). *Diagnostic Evaluation of Language Variation (DELV) Technical Manual*. San Antonio: Harcourt Inc.

Shatz, M., Wellman, H., & Silber, S. (1983). The acquisition of mental verbs: a systematic investigation of first references to mental state. *Cognition, 14*, 301-321.

Shepard, L., Kagan, S. L., & Wurtz, E. (Eds). (1998). *Principles and recommendations for early childhood assessments*. Washington, DC: national Education Goals Pannel.

Snow, C. E., & Dickinson, D. K. (1990). Social sources of narrative skills at home and at school. *First Language, 10*, 87-103.

Stockman, I. J. (1996). The promise and pitfalls of language sample analysis as an assessment tool for linguistic minority children. *Language, Speech, and Hearing Services in Schools, 27(4)*, 355–66.

Stockman, I. J. (2000). The new Peabody Picture Vocabulary Test–III: An illusion of unbiased assessment? *Language, Speech, and Hearing in the Schools, 31(4)*, 340–353.

Stockman, I. J., & Vaughn-Cooke, F. (1984). Semantic categories in the language of working class black children. In C. E. Johnson & C. L. Thew (Eds.), *Proceedings of the Second International Child Language Conference, Vol. 1*, 312–327.

Stockman, I. J., & Vaughn-Cooke, F. (1986). Implications of semantic category research for language assessment of nonstandard speakers. *Topics in Language Disorders, 6(4)*, 15–25.

Tabors, P. O., Roach, K. A., & Snow, C. E. (2002). How language and literacy environment: Final results. In D.K. Dickinson and P.O. Tabors (Eds.), *Beginning literacy with language: Young children learning at home and learning at school.* (pp. 111-138).

Thompson, R. A. (1994). Emotion regulation: A theme in search of a definition. In N. Fox (Ed.). *The development of emotion regulation. Monographs of the Society for Research in Child Development, 59*, 25-52.

Washington, J. A., & Craig, H. K. (1991). Performances of low-income, African American preschool and kindergarten children on the Peabody Picture Vocabulary Test—Revised. *Language, Speech and Hearing Services in Schools, 23(4)*, 329–333.

Waters, E., & Deane, K. (1985). Defining and assessing individual differences in attachment relations: Q-methodology and the organization of behavior in infancy and early childhood. In I. Bretherton & E. Waters (Eds.), *Growing points of attachment theory and research. Monographs of the Society for Research in Child Development, 50* (pp. 217-2555). Hillsdale, NJ: Erlbaum.

Waxman, S. R., & Hatch, T. (1992). Beyond the basics: Preschool children label objects flexibly at multiple hierarchical levels. *Journal of Child Language, 19(1)*, 153–166.

Wyatt, T. (Chair) (2003). *The assessment of African American children: An update on Larry P. California Speech-Language and Hearing Association Task Force*, January 7, 2003.

APPENDIX

Temple University Forum on Preschool Assessment
January 30-31, 2003

AGENDA

Thursday, Jan. 30: Part A: What Should Be Assessed and Why?

8:00 a.m. - 9:00 a.m. Continental Breakfast – Room 220
9:00 a.m. - 10:30 a.m. Orientation and Background – Room 220
Discussion facilitated by Kathy Hirsh-Pasek

10:30 a.m. - 12:30 p.m. Parallel Workshop Sessions
1. Language/Literacy
Discussion facilitated by Kathy Hirsh-Pasek, Room 205
2. Numeracy/Spatial Ability
Discussion facilitated by Nora Newcombe, Room 305
3. Social-emotional Development
Discussion facilitated by Marsha Weinraub, Room 604

12:30 p.m. - 2:00 p.m. Lunch – Room 220
2:00 p.m. - 3:30 p.m. Workshop Session Reports – Room 220
3:30 p.m. - 4:00 p.m. Afternoon Break
4:00 p.m. - 5:00 p.m. Looking Toward Assessment – Room 220

Friday, Jan. 31: Part B: What Assessment Tools Do We Have? What Do We Need?

8:00 a.m. - 8:30 a.m. Continental Breakfast – Room 220

8:30 a.m. - 10:30 a.m. Parallel Workshop Sessions – Room 220
1. Language/Literacy
Discussion facilitated by Kathy Hirsh-Pasek, Room 205
2. Numeracy/Spatial Ability
Discussion facilitated by Nora Newcombe, Room 305
3. Social-emotional Development
Discussion facilitated by Marsha Weinraub, Room 604

10:30 a.m. - 11:00 a.m. Break
11:00 a.m. - 12:00 p.m. Workshop Session Reports – Room 220
12:00 p.m. - 1:00 p.m. Lunch – Room 220

Part C: Next Steps?

1 a.m. - 3 p.m. Discussion facilitated by Nora Newcombe – Room 220

Center for Improving Resources in Children's Lives Temple University

CIRCL is a center where cutting edge foundational research is translated into real world applications. A major gap exists between basic research on children's development and the public policies and practices that affect children and their families. CIRCL's mission is to bridge this gap by bringing together researchers, practitioners, policymakers and parents so that evidence-based understanding of early child development can be achieved.

For more information, contact:

Anita T. Kochanoff, Ph.D.
Visiting Assistant Professor
Department of Psychology
Programs Director, CIRCL

Temple University
563 Weiss Hall
1701 N. 13th Street
Philadelphia, PA 19122

Phone: 215-204-5744
Fax: 215-204-2715
Email: anita.kochanoff@temple.edu
www.temple.edu/CIRCL

For more information, contact:

Anita T. Kochanoff, Ph.D.
Visiting Assistant Professor
Department of Psychology
Programs Director, CIRCL

Temple University
563 Weiss Hall
1701 N. 13th Street
Philadelphia, PA 19122

Phone: 215-204-5744
Fax: 215-204-2715
Email: anita.kochanoff@temple.edu

www.temple.edu/CIRCL

Temple University Forum on Preschool Assessment sponsored by:
Department of Psychology
Center for Public Policy
College of Liberal Arts
Office of the Vice President for Research and Graduate Studies

CIRCL

Translating research into reality

Temple University