TO: Katie Fallin

FROM: Sally Atkins-Burnett, Maggie Caspe, Cheri Vogel, Susan Sprachman, and John Love

DATE: 10/19/2007

UPCOS-195

SUBJECT: Fall Child Assessment Battery Decisions

The overall goal of the First 5 LA Universal Preschool Child Outcomes Study (UPCOS) is learning how children enrolled in the Los Angeles Universal Preschool (LAUP) and Power of Preschool (PoP) programs are faring. In this two-phase study, Mathematica Policy Research, Inc. (MPR), along with its partners (Juárez & Associates; American Institutes for Research; Berkeley Evaluation and Assessment Research Center; and Deanna Gomby, a senior consultant), is working with First 5 LA, LAUP, and their Research Advisory Committee (RAC) to design and implement the study.

Phase 1 Pilot Test

In Phase 1 of the study, we examined the feasibility, reliability, and validity of various child development measures in the large and culturally and linguistically diverse population of children served by the LAUP programs in Los Angeles County. The measures selected for use in the pilot test had the following features:

- Collectively covered all the domains of children’s early development and learning that comprise school readiness
- Demonstrated evidence of use with diverse populations (including linguistic, racial/ethnic, and ability diversity)
- Demonstrated evidence of reliability and validity
- Balanced the “tried and true” with the new
- Were sensitive to interventions
- Were age appropriate for 4-year-olds
- Were available (as much as possible) in both Spanish and English
The pilot sample included 417 children from diverse backgrounds who attended the LAUP program during the 2006–2007 program year. All children were given language screening, receptive and expressive language, and executive functioning assessments. Children in each class were then randomly assigned to receive either mathematical or literacy assessments (rapid letter naming and phonological awareness). The data from the pilot test were analyzed for appropriateness of use with children in the LAUP programs. The results of these analyses were discussed with the RAC and with First 5 LA and LAUP to make decisions about the Phase 2 child assessment battery to be administered in fall 2007 and spring 2008.

The results of these decisions are displayed in Table 1, which provides information on the child assessment measures that are being used in the Phase 2 data collection. Additionally, the table outlines the general purpose of the measure; whether the measure was retained from the spring or is a replacement measure; and, when appropriate, the rationale for the replacement or nonretention. Appendix A describes each measure and summarizes the results of the analysis in greater detail.

Table 1. Fall 2007 Child Assessment Measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>Purpose</th>
<th>Retained from Pilot</th>
<th>Replacement</th>
<th>Rationale for Nonretention or Replacement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spanish Pre-LAS: Tio Simon Dice</td>
<td>Language Screener: Spanish receptive proficiency</td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Spanish Pre-LAS: Exposición de Arte</td>
<td>Language Screener: Spanish expressive proficiency</td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>English Pre-LAS: Simon Says</td>
<td>Language Screener: English receptive proficiency</td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>English Pre-LAS: Art Show</td>
<td>Language Screener: English expressive proficiency</td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Peabody Picture Vocabulary Test (PPVT-4) and Test de Vocabulario en Imágenes Peabody (TVIP)</td>
<td>Language Development: Receptive vocabulary</td>
<td>No</td>
<td>No</td>
<td>Analysis of the PPVT-4 identified misfitting items and items with differential item functioning between English speakers and Spanish speakers.</td>
</tr>
<tr>
<td>Receptive One-Word Picture Vocabulary Test (ROWPVT)*</td>
<td>Language Development: Receptive vocabulary</td>
<td>Yes</td>
<td>Replaces the PPVT-4</td>
<td></td>
</tr>
<tr>
<td>Measure</td>
<td>Purpose</td>
<td>Retained from Pilot</td>
<td>Replacement</td>
<td>Rationale for Nonretention or Replacement</td>
</tr>
<tr>
<td>--------------------------------------------------------------</td>
<td>---------------------------------------</td>
<td>---------------------</td>
<td>-------------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>Expressive One-Word Picture Vocabulary Test (EOWPVT)(^a)</td>
<td>Language Development: Expressive vocabulary</td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Woodcock-Johnson Applied Problems and Woodcock-Muñoz Bateria</td>
<td>Cognitive Development: Math concepts</td>
<td>No</td>
<td>No</td>
<td>The English measure had acceptable reliability (.84), and only one item with differential item functioning (DIF). However, we dropped this measure to reduce the burden on children. The ECLS-B provides a measure of mathematics that can be conceptually scored and provides information on more extensive mathematical content.</td>
</tr>
<tr>
<td>ECLS–B Math(^a)</td>
<td>Cognitive Development: Math concepts</td>
<td>Yes</td>
<td>No</td>
<td>Based on the need to conserve time and an analysis of the data, we will use only the ECLS-B Math section, eliminating the Woodcock-Johnson Applied Problems test and its Spanish counterpart, the Woodcock-Muñoz Bateria.</td>
</tr>
<tr>
<td>Height and Weight</td>
<td>Physical Development</td>
<td>Yes</td>
<td></td>
<td>We added this at the RAC’s recommendation to provide a measure of children’s physical development.</td>
</tr>
<tr>
<td>Executive Functioning</td>
<td></td>
<td></td>
<td></td>
<td>Note: The Draw a Circle task was dropped from the executive functioning battery to reduce the burden on children. The addition of the Woodcock-Johnson and Woodcock-Muñoz Spelling provides a measure of visual motor</td>
</tr>
</tbody>
</table>
MEMO TO: Katie Fallin  
FROM: Sally Atkins-Burnett, Maggie Caspe, Cheri Vogel, Susan Sprachman, and John Love  
DATE: 10/19/2007  
PAGE: 4

<table>
<thead>
<tr>
<th>Measure</th>
<th>Purpose</th>
<th>Retained from Pilot</th>
<th>Replacement</th>
<th>Rationale for Nonretention or Replacement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rapid Letter Naming&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Literacy Development: Letter identification</td>
<td>Yes</td>
<td>No</td>
<td>coordination.</td>
</tr>
<tr>
<td>Development of English Literacy in Spanish-Speaking Children (DELLS) : Phonological Awareness Test (English and Spanish versions)</td>
<td>Literacy Development: Phonological awareness</td>
<td>No</td>
<td>No</td>
<td>For Spanish speakers, we found poor reliability and floor problems. The English sample had ceiling problems.</td>
</tr>
<tr>
<td>Woodcock-Johnson III—Test 7 Test Spelling and Woodcock-Muñoz III—Test 7 Ortografiá</td>
<td>Literacy Development and Fine Motor Skills: Writing and spelling/Ability to take dictation/Fine motor control</td>
<td>Yes</td>
<td></td>
<td>The English and Spanish version of this measure replaced the DELLS. The current measure is standardized, reliable, and valid with national norms.</td>
</tr>
<tr>
<td>SPRING ONLY Environmental Print Awareness</td>
<td>Literacy Development: Ability to identify words in the environment</td>
<td>Yes (Spring 2008 only)</td>
<td></td>
<td>The RAC recommended that we collect additional information about literacy. This will be used in the spring to measure children’s ability to recognize print common in the classroom and community.</td>
</tr>
<tr>
<td>Leiter-R Examiner Rating Scales</td>
<td>Cognitive Development: Assessor rating of children’s attention, activity level, and sociability</td>
<td>Yes</td>
<td>No</td>
<td>Measure was reduced from four subscales to three (eliminating Self-Regulation, which was difficult for assessors to observe within the context of this assessment).</td>
</tr>
</tbody>
</table>

**Parent Measures**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Purpose</th>
<th>Retained from Pilot</th>
<th>Replacement</th>
<th>Rationale for Nonretention or Replacement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health Ratings</td>
<td>LA County Health Survey items</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preschool Kindergarten Behavior Scales–2 (PKBS-2)</td>
<td>Rating positive social skills</td>
<td>Yes</td>
<td></td>
<td>This measure was adapted based on pilot data.</td>
</tr>
</tbody>
</table>

**Teacher Measures**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Purpose</th>
<th>Retained from Pilot</th>
<th>Replacement</th>
<th>Rationale for Nonretention or Replacement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preschool Kindergarten Behavior Scales–2 (PKBS-2)</td>
<td>Rating positive social skills</td>
<td>Yes</td>
<td></td>
<td>In response to teacher concerns about the number of problem behaviors, we will use a shorter 10-item scale from</td>
</tr>
<tr>
<td>Social Skills Rating System (SSRS)</td>
<td>Rating problem behaviors</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Prepared by Mathematica Policy Research (UPCOS-195)*
Measure | Purpose | Retained from Pilot | Replacement | Rationale for Nonretention or Replacement
--- | --- | --- | --- | ---
Desired Results Developmental Profile–Revised (DRDP-R) | Comprehensive rating of child development | Yes | Only PoP programs will complete the DRDP-R.

\*Denotes conceptual scoring.

**Language Routing**

An innovation of the Phase 1 pilot test was the development of a language routing protocol that gives children multiple chances to be routed into their dominant language (English or Spanish). We used three sources of information: (1) parent report of children’s language use in the home and community, (2) performance on two subtests from the Oral Language Development Scale (OLS) of the Pre-LAS 2000 (Duncan and DeAvila 2002), and (3) a conceptually scored receptive vocabulary test.

The consent form asks parents to report on their home language, the language families use in speaking to the child, the language the child uses with the parents, and the language the child uses with other children (Table 2). Parent reports of children’s language guide the use of language screeners (the Simon Says and Art Show tasks from the Pre-LAS 2000). Children who speak Spanish primarily are administered the Spanish warm-up portion of the screening and then are administered the English version of the screening (with the items ordered differently). In combination, the parent report and English Pre-LAS help determine which children should receive the rest of the assessment battery presented first in English or Spanish. Finally, the Receptive One-Word Picture Vocabulary Tests (Brownell 2001) is administered. Children with a higher proportion of items correct in English are administered probes in English first (e.g., “point to ball”), whereas children with a higher proportion of items correct in Spanish are administered probes in Spanish first (e.g., “señala pelota”).

Prepared by Mathematica Policy Research (UPCOS-195)
Table 2. Parents’ Report of Children’s Language Context

<table>
<thead>
<tr>
<th>CHECK ONE BOX PER LINE</th>
<th>Only English</th>
<th>Mostly English but sometimes my primary language</th>
<th>Both languages about equally</th>
<th>Mostly my primary language but some English</th>
<th>Only my primary language</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. What language do you speak most often with your child?</td>
<td>□ 5</td>
<td>□ 4</td>
<td>□ 3</td>
<td>□ 2</td>
<td>□ 1</td>
</tr>
<tr>
<td>2. What language does your child speak most often with you?</td>
<td>□ 5</td>
<td>□ 4</td>
<td>□ 3</td>
<td>□ 2</td>
<td>□ 1</td>
</tr>
<tr>
<td>3. What language does your child speak most often with other children?</td>
<td>□ 5</td>
<td>□ 4</td>
<td>□ 3</td>
<td>□ 2</td>
<td>□ 1</td>
</tr>
</tbody>
</table>

In addition to the routing mechanism, the assessment allows “conceptual scoring,” meaning that children answering correctly in either English or Spanish get credit for that response, regardless of the language in with the question was asked. In the ROWPVT, the prompt is first administered in the language that was determined to be the child’s dominant language for the assessment, and then it is repeated in the other language if the child responds incorrectly.

It is important to note that children who speak neither English nor Spanish as their dominant language and score fewer than 15 items correct on the ROWPVT skip to the height and weight section and are then routed out of the evaluation. Figure 1 illustrates the entire routing pathway.
MEMO TO: Katie Fallin  
FROM: Sally Atkins-Burnett, Maggie Caspe, Cheri Vogel, Susan Sprachman, and John Love  
DATE: 10/19/2007  
PAGE: 7

Figure 1. Language Routing Pathway

Prepared by Mathematica Policy Research (UPCOS-195)
Appendix A

This appendix presents in greater detail the analyses we did using the spring Phase 1 pilot sample. It provides supporting information for the decisions made for the final battery as described in the main document. These decisions balanced the psychometric properties of the instruments, the time required to administer them, and the preferences of First 5 LA and the Research Advisory Committee (RAC). Rather than being organized around the order of administration as in the body of this document, this appendix first describes direct and then indirect measurements of children’s development.

Direct Measures of Children’s Development

Development of English Literacy in Spanish-Speaking Children—Phonological Awareness Test I (DELLS Phonological Awareness I; Harvard University and Center for Applied Linguistics 2002). Though this relatively new measure had limited reliability data, it was included for piloting based on the importance of phonological skills and the strength of having been evaluated in English and in Spanish. Each version of the measure included 18 items that ask children to identify words that rhyme or produce a rhyme and to identify words that begin with the same sound (alliteration). The English version was administered to English speakers and the Spanish version to children for whom Spanish was their first language. However, the reliability of the measure with the Spanish sample was poor (KR-20 = .53). The English version had ceiling problems (15 percent scored correct on all items) and the Spanish version had floor problems (9 percent did not get any items correct). The Spanish rhyming items were much more difficult than the English rhymes. The most difficult English rhyming items were multisyllabic. All the Spanish rhyming items were multisyllabic. As a result of low performance on rhyme recognition, children who did the Spanish assessment were routed out of rhyme production. Given the problems with this measure, a decision was made to drop it from the battery.

Rapid Letter Naming Task. A criterion reference measure of the number of letters that a child could name was used in the pilot study. Letter naming (Moats 1998; Snow, Burns, and Griffin 1998) and, in particular, rapid letter naming (O’Connor and Jenkins 1999; Rouse and Fantuzzo 2006; Uhry 2002) are among the strongest predictors of later success in school. We piloted a rapid letter-naming task that was very reliable (KR-20 = .97) and provided information about which letters children were able to name easily and rapidly. As expected, children were able to name more uppercase than lowercase letters.

Two different forms were developed with some overlap of items. Each form had some uppercase and some lowercase items. If an item was uppercase on one form, it was lowercase on the other form. All children were asked about either the uppercase or lowercase form of every
letter of the alphabet. Very common letters (such as vowels) were presented as both uppercase and lowercase on both forms to help provide overlap in scaling the measure. As in the pilot, children in each classroom will be randomly assigned to one of these forms. Each child will be asked to name only 30 uppercase or lowercase letters and the examiner will move to the next letter after three seconds. Both English and Spanish names will be accepted. The 30 letters can be administered in less than two minutes.

In the spring pilot, 4 percent of the children were able to name all 30 letters administered to them and many (about 35 percent) named more than 25 letters. The RAC pointed to the need to collect more data about literacy than just letter naming. Consideration was given to including a norm-referenced measure of concepts of print or spelling (written literacy). Standardized measures of concepts of print are not available in both English and Spanish. Therefore, we recommended that environmental print words be added to the letter-naming task and the Woodcock-Johnson–III Spelling subtest be used. The WJ-III Spelling subtest is a nationally standardized assessment and has been used in FACES 2006 as well as previous FACES and has established reliability and validity evidence.

The Environmental Print Awareness words will be added to the end of the assessment and will include both Spanish and English words. Each word will be presented separately in English and Spanish. A list of potential words was generated and feedback was obtained from the LAUP coaches regarding the words children are most likely to observe in their classrooms. In addition, four common sight words (is, me, es, mi) that children would see if teachers model sentence writing and take children’s dictation were added to the list of words. The test words were selected based on the frequency with which they would occur in English and Spanish sentences. Children will be credited with the correct Spanish or English translations of the words or with partial credit if they name a word that begins with the same sound but is not the test word (e.g., if for “table” they say “top,” they would receive partial credit for recognizing the sound of the initial consonant).

Woodcock-Johnson Psycho-Educational Battery (and Batería III Woodcock-Muñoz) Third Edition (WJ-III) (Woodcock, McGrew and Mather 2001). We administered the Applied Problems subtest from the English and Spanish versions of this battery. Children could respond in either language. This subtest measures a child’s skill in analyzing and solving practical problems in mathematics. The reliability of the scales with our sample was acceptable (.84). The bivariate correlation with the receptive language measure (PPVT-4; r = .64) was as strong as the correlation with the ECLS Mathematics measure (r = .63), indicating a strong influence of language on the WJ-III Applied Problems subtest. It also showed a slightly stronger relationship to measures of social skills than did the ECLS Mathematics measure. Only one item showed differential item functioning (DIF). This test is the most frequently used test of early childhood mathematics in large-scale studies. Although we had initially recommended retaining it to provide comparisons to other samples and studies, we removed it to lessen the burden for children.

Prepared by Mathematica Policy Research (UPCOS-195)
Early Childhood Longitudinal Study Mathematics. We supplemented the WJ-III Applied Problems subtest with items from the ECLS-B and ECLS-K mathematics assessments that tap additional areas of mathematical skills, such as spatial abilities and measurement. Items in Spanish and English were identical, and children could respond in either language. The reliability of the scale was adequate (Rasch scale reliability = .80), and no DIF was identified. These items were developed for this age group and target the range of mathematical concepts typically developing among 4-year-old children. We plan to anchor on the national item difficulties to provide a national comparison in the spring.

Language screener and language of administration. The language screener used the Pre-LAS Simon Says and Art Show as well as parent reports on the language(s) a child speaks at home and to other children to determine whether a child should receive the English or the Spanish version of the child assessment battery or not be tested at all. Figure 1 illustrates the flow of the screening and how it will inform the language of administration. We used receiver operating curve (ROC) analysis with the spring pilot data to determine the optimal cutoff for ensuring that children are fairly assessed in the appropriate language.

Whenever we are assessing mathematics, letter identification, and vocabulary, we will use conceptual scoring (Bedore et al. 2005), allowing the child to respond in the language that he or she wishes to use. For Spanish and English, we will credit the child for concepts that are correct regardless of the language used. We will note the language in which the child responds. Our assessors will be fluent in Spanish and English. Assessments are not available in other languages, so we are limited in our ability to extend this to other groups. Per the recommendation of the RAC to also obtain a measure of English-language proficiency, the receptive vocabulary will be assessed in the spring only in English (see ROWPVT below) although it is administered conceptually in the fall.

Peabody Picture Vocabulary Test–Fourth Edition (PPVT-4). The PPVT-4 (Dunn and Dunn 2006) is the measure of receptive vocabulary piloted in the spring with children who passed the English screener. The Spanish-language version of the PPVT, the Test de Vocabulario en Imagenes Peabody (TVIP; Dunn et al. 1986), was used with children whose primary language was Spanish. The TVIP was also administered to Spanish-speaking children who were able to complete the assessments in English in order to describe their overall receptive vocabulary proficiency and examine potential cross-language transfer of language abilities. Preliminary analysis of the PPVT-4 with our sample identified seven misfitting items and seven items with potential DIF (four favoring English speakers and three favoring Spanish speakers). Scale reliability was high (.97); however, based on concerns about potential misfit and DIF, as well as the difficulties in comparing performance on the PPVT-4 with the dated norms on the TVIP, we decided not to administer the PPVT-4 or the TVIP.
Instead of the PPVT-4/TVIP, we will administer the Receptive One-Word Picture Vocabulary Test: English and Spanish–Bilingual Edition (ROWPVT; Brownell 2001) as our measure of receptive vocabulary. The ROWPVT can be administered in English, Spanish, or as a bilingual measure. In the spring of year 2, we will administer the ROWPVT in English only as a measure of English-language vocabulary. The ROWPVT corresponds to the EOWPVT and provides a stronger basis of comparison between the receptive vocabulary and expressive vocabulary since they were standardized on the same sample.

Expressive One-Word Picture Vocabulary Test: English and Spanish–Bilingual Edition (EOWPVT; EOWPVT-SBE; Brownell 2000). The EOWPVT-SBE was administered to all children and scored conceptually according to the standardized scoring. The reliability was high (.96). The preliminary analysis identified four misfitting items and three items that favored English speakers in this smaller sample.

We considered the inclusion of a sentence repetition task in order to get a measure of children’s language beyond vocabulary. We investigated the Woodcock-Muñoz Language Battery. This task is not available in the English version. We could not locate a standardized preschool sentence imitation task available in both English and Spanish.

Executive Functioning Tasks: Pencil Tapping, Draw a Circle, Balance Beam (Walk a Line Slowly). These “cool” executive functioning tasks were easy to administer and provide information about children’s executive functioning as well as their motor skills. We found variation in scores and are continuing to examine the validity of these scores. We recommend that these tasks be administered again in the fall. However, since the addition of the WJ-III Spelling subtest and measurement of height and weight added burden, we decided to administer only two of the three tasks. The WJ-III Spelling subtest provides a measure of visual motor skills as well as written literacy, so we excluded the Draw a Circle task.

Although cool tasks such as the ones we piloted measure a child’s ability to regulate impulses and to exercise “effortful control,” “hot” tasks measure emotional regulation and typically involve food, gifts, or frustration (Kochanska and Murray 2002). The RAC had a question regarding the potential inclusion of a hot task. We discussed the problems with administration of these tasks (e.g., use of food and differences in what is emotionally arousing to children particularly when used with diverse cultures) and decided to continue with just the cool tasks.

Height and Weight. Based on the recommendations of the RAC, we will weigh the children and measure their height in the fall and spring.
Indirect Measures of Children’s Development

**ECLS-K Approaches to Learning.** Teachers rated each child on the six items that comprise the Approaches to Learning Scale from the Early Childhood Longitudinal Study—Kindergarten Class or 1998–1999 (ECLS-K) study. The reliability of the scale in our sample was strong (KR-20 = .91), although there was a ceiling problem (27 percent at maximum score). The bivariate correlations with measure of mathematics (r = .35–.37) and language (r = .21–.24) were significant. The teachers will complete this measure in the fall.

**Leiter-R Examiner Rating Scales.** At the conclusion of the direct assessment, assessors will rate children on subtests from the Leiter-R Examiner Rating Scales to obtain a more comprehensive picture of each child’s attention, activity level, and sociability. The items on these subscales are drawn from the cognitive/social scale. In the pilot, we included the items from the Leiter-R subscale for Self-Regulation, but the items are not aligned with our testing protocol, so we are eliminating that subscale. We have a direct measure of inhibition in the executive functioning tasks (Roid and Miller 1997).

**Parent Ratings of Health.** The RAC, recommended that we use scales from the health survey that Frances Glascoe developed in order to have a standardized measure of health and physical well-being. Some of the questions in the parent interview are drawn from the LA County Health Survey to provide a comparison with the overall population in LA County. (First 5 LA reported that the LA County Health Survey adapted the Glascoe survey.) First 5 LA expressed a preference for being able to compare study data to local data.

**Rating of Social-Emotional Development.** Both parents and teachers will rate children’s behavior and social-emotional development. The teachers will rate children’s social-emotional competence using the positive social scales from the *Preschool Kindergarten Behavior Scales*–2 (PKBS-2; Merrell 2002) and rate problem behaviors with the *Social Skills Rating System* (SSRS; Gresham and Elliott 1990). The teachers were positive about the PKBS-2 social scales, and some teachers commented that they captured more typical behaviors and were more closely aligned with common preschool behaviors. Teachers were concerned about the high number of problem behaviors included. To reduce burden and respond to teacher concerns, we will use the shorter 10-item scale from the SSRS.

Teachers preferred a frequency rating (“how often” rather than “how true”) and expressed a preference for more points on a scale (four- to five-point scales rather than three-point scales). Although these changes deviate from the normed version, we decided to make this adjustment. With this caveat, use of the standardized scale will allow a comparison to national norms provided by the publishers of these instruments, so the teacher ratings will use the published standard version. The parents will complete an adapted form of the PKBS-2. Items were adapted and the measures abbreviated based on each item’s contribution to the scale (factor loadings) and

*Prepared by Mathematica Policy Research (UPCOS-195)*
Desired Results Developmental Profile–Revised (DRDP-R). The DRDP-R is implemented statewide as the required assessment tool for all care programs funded by the California Department of Education, Child Development Division. This observational assessment is designed to measure children’s development in all areas of child outcomes: social skills, language, literacy, mathematics, self-regulation, approaches to learning, cognitive and general knowledge, health, and physical development. All the scales within the DRDP-R showed strong reliability (.89 to .96). A multidimensional Rasch model was used to estimate scores, with the scale projected on a common axis for the sake of comparability of dimensions. This method of calibration combined with the shared method variance and an adjustment of correlations for the measurement error increased the inter-factor correlations (r = .82–.99). Separate calibrations of the subscales indicated lower correlations, although there appears to be a strong higher order factor (see the separate report on the DRDP-R for additional information about the analysis of the DRDP-R). The language and literacy dimension shows evidence of convergent and divergent validity with the direct measures. In the area of language, the items on the DRDP-R extend beyond vocabulary and assess comprehension of increasingly complex language (e.g., attributes and time and causality relationships), ability to follow instructions, conversational skills, and expressive language (including semantic and syntactic complexity).

The other dimensions on the DRDP-R are also moderately correlated with direct measures of vocabulary. This suggests that the teachers are strongly influenced by children’s verbal skills when rating them on all the domains. Much of what teachers know about children’s skills and knowledge is based on what the child says.
MEMO TO: Katie Fallin  
FROM: Sally Atkins-Burnett, Maggie Caspe, Cheri Vogel, Susan Sprachman, and John Love  
DATE: 10/19/2007  
PAGE: 14

References


Prepared by Mathematica Policy Research (UPCOS-195)

