Predicting Poor Achievement in Early Grade School Using Kindergarten Scores on Simple Cognitive Tasks

Marcia Strong Scott and Christine F. Delgado
University of Miami

Abstract: In this study, kindergarten children’s scores on nine cognitive tasks were evaluated as potential predictors of poor achievement in first grade. A set of five tasks successfully identified 83% of children who were poor readers in first grade, while just three tasks identified 72% of children who were poor achievers in math in first grade. There were a large number of false positives in the predictions to first grade achievement, but some of these children were poor achievers in subsequent grades. Many of the children who were poor achievers in reading or math were actually poor achievers in both academic areas. How such screening results might be used was discussed.

Scott and her colleagues have been working to develop more effective screening tests to identify young children with mild learning problems who will require special education services in early grade school (e.g., Scott & Fletcher, 2001; Scott, Fletcher, Jean-Francois, Urbano, & Sanchez, 1998). The use of such screening measures is based on the assumption that one can improve the cognitive performance of young children through early intervention (Salvia & Ysseldyke, 1991). Research results have shown that participation in such interventions is most effective if begun at an early age (Shonkoff & Hauser-Cram, 1987), is associated with a reduced need for subsequent special education (Lazar & Darlington, 1982), and is particularly effective for those children with only mild delays (Dunst, Trivette, & Cross, 1986). Such results support the need to develop more effective screening instruments to identify children with mild learning problems at an early age so educational interventions can begin when they are most efficacious.

The screening batteries were composed of simple cognitive tasks, but tasks that required the active engagement of the children’s cognitive processing skills (Greenfield & Scott, 1985). For example, children might be shown a page on which four people differing on several dimensions were displayed and be asked to tell how people differ from one another (Scott, Deuel, Urbano, Fletcher, & Torres, 1998). To respond correctly to tasks such as this one, children must activate their cognitive processing capabilities and analyze the problem, often tapping into their long-term memory. A simple automatic associative response will not do. The utilization of their cognitive processing capabilities was expected to make overt, their not-yet-identified mild cognitive deficiencies.

The target groups with mild learning problems used to evaluate potential test items consisted of young children with either an educable mentally handicapped (EMH) or specific learning disabled (SLD) classification (Gresham, MacMillan, & Bocian, 1996; Holtzman & Wilkinson, 1991).

Studies designed to select and assess the concurrent validity of screening sets, using
matched samples of children with and without mild learning problems, typically were associated with high levels of both sensitivity and specificity, and low percentages of overreferrals (e.g., Scott, Deuel, et al., 1998; Scott, Fletcher, & Martell, 2000). When the preschool screening set was cross validated with an unmatched sample consisting of a large number of children without learning problems and many fewer children with learning problems (Scott, Tu, & Fletcher, 2003), sensitivity and specificity values remained at levels greater than 80%, the minimum levels of classification accuracy recommended by Meisels (cited in Lichtenstein & Ireton, 1991). However, for the first time, the percentage of overreferrals was high, although not much above the suggested percentage of normally achieving children needed to be referred in order to identify children from low base rate groups (Lichtenstein & Ireton, 1984).

In addition to evaluating potential test items using concurrent educational classifications, the best items to constitute a screening set must also be selected and evaluated in terms of their long-term predictive accuracy (Rafoth, 1997) so that examiners can judge how well initial screening results can predict future diagnostic/placement decisions (Lenkarski, Singer, Peters, & McIntosh, 2001). Two predictive studies were conducted to select the best items for a preschool (Scott & Delgado, 2003) and a kindergarten (Scott, Delgado, Tu, & Fletcher, 2005) screening test. In both studies, the full screening battery was administered to a very large sample of children without any learning problems. Risk for a special education placement of EMH or SLD in third grade was the outcome measure. Risk for a special education placement of EMH or SLD in third grade was the outcome measure. In both studies, there was a large percentage of false positives which, combined with the low prevalence target group, resulted in a high percentage of overreferrals. However, for these studies, achievement performance data on the reading comprehension, math application and math computation subtests of the Stanford Achievement Test were available and we were able to determine the third grade achievement performance of these false positives, looking to see if any were exhibiting problems in academic achievement.

We typically evaluate potential test items in terms of their contribution to the classification accuracy achieved when differentiating two different educational groups. This method of evaluation is the recommended method for assessing any screening test (e.g., Lichtenstein, 1981; Lichtenstein & Ireton, 1984, 1991). Differentiating between children who were, or were not poor achievers was anticipated to be an even more difficult differentiation to make than the differentiation of children in, or not in special education. We chose, therefore, to define poor achievement as having an achievement score that fell at, or less than the 15th percentile, rather than the more typical definition of poor or low achievement which is earning an achievement score that falls below the 25th percentile (e.g., Algoyzine & Ysseldyke, 1983; Shaywitz, Fletcher, Holahan, & Shaywitz, 1992; Shinn, Ysseldyke, Deno, & Tindal, 1986; Ysseldyke, Algoyzine, Shinn, & McGue, 1982).

This definition of poor achievement was used to evaluate children in the preschool (Scott & Delgado, 2003) and kindergarten (Scott et al., 2005) studies who earned screening scores below the cut for a prediction of risk for a special education placement in third grade but who, in third grade, were in regular, not special education (false positives) and who could be classified as a poor achiever or not. We found that 34% of the preschool sample and 57% of the kindergarten sample who met these criteria were poor achievers in third grade. In effect, a good percentage of the false positives in both studies using the risk-for-a-special-education-placement criterion actually were at risk for another type of mild learning problem; poor academic achievement.

These results led us to consider whether children’s performance on a select, albeit different subset of cognitive tasks from the same battery might be used to identify kindergarten children at risk for poor achievement rather than risk for a special education placement. Cognition is, after all, an area related to later school achievement (Funk, Sturner, & Green, 1986) and cognitive capabilities assessed prior to, or at the beginning of kindergarten have been shown to relate to both reading and math performance (e.g., Mantzicopoulos & Morrison, 1994; Swanson, 1994; Tramontana, Hooper, & Selzer, 1988). For this study, as in previous screening studies, individual tasks were evaluated in terms of
their contribution to classification accuracy. In this instance, however, it was classification accuracy associated with the identification of those kindergarten children who were, or were not at risk for poor achievement in first grade.

**Method**

**Participants**

**Kindergarten sample.** Four hundred and fifty-nine kindergarten children were administered a screening battery consisting of nine cognitive tasks in the 1993-94 school year. This sample was composed of every child from the targeted kindergarten classes in 35 participating public schools for whom an informed consent was obtained. Informed consent included permission to access the child’s educational status and SAT scores through third grade. All six regions of the Miami/Dade County Public School (M/DCPS) system were represented. According to parental designations, 39% of the sample were classified as Black, 34.5% as Hispanic, and 26.5% as White. These three racial/ethnic classifications were the only ones used by the school system at that time. Fifty-nine percent of the sample was male and the mean chronological age was 72.0 months ($SD = 4.2$).

**First grade regular education sample.** Children were tracked into grade school based on their unique school identification number. Three hundred and eighty children from the kindergarten sample were located in first grade in the M/DCPS data base. Children not in this sample were either not in the M/DCPS first grade data base in the year after kindergarten or they were in special, not regular education.

The distribution of race/ethnicity in the first grade sample was nearly identical to that for the kindergarten sample: 41% were designated Black, 33% Hispanic and 26% White. Clearly, the majority of children in both samples were from minority groups. These primarily minority samples mimicked the total school population which was approximately 83% minority in the year the kindergarten children were administered the screening battery.

Fifty-seven percent of the first grade sample were males which was nearly identical to the percentage of males in the kindergarten sample. The mean chronological age of the first grade sample when they were in kindergarten, was the same ($M = 72.0, SD = 4.2$) as the mean chronological age of the 459 children in the kindergarten sample ($M = 72.0, SD = 4.2$).

**Testing Language**

If a child lived in a home in which a language other than English was spoken, that child was tested in the language in which the child was most proficient. Proficiency was determined through a consideration of teacher nomination, the child’s choice and the child’s conversational ability. Spanish was the only other language used for testing. If the examiner was not certain that the child understood the instructions in the language employed, the examiner administered instructions in both languages. Examiners fluent in Spanish translated instructions into that language for those children who needed to be tested in Spanish or both languages. Eighty-one percent of the 158 Hispanic children in the kindergarten sample chose to be tested in English and the others were tested in Spanish (15%) or both languages (4%). A similar distribution of testing language was found for the 125 Hispanic children who were in the first grade sample. The screening battery had been administered in English to 86% of this sample, in Spanish to 12% and in both English and Spanish to just 2%. Responses in Spanish were accepted from any child regardless of testing language.

**Screening Battery Contents and Administration**

The screening battery consisted of nine tasks. Test items included colored photographs of meaningful pictures, or in one case colored dots. These stimuli were placed on 35.6 cm wide by 21.6 cm high white paper pages, which were placed in a four-hole legal size black binder, 36.8 cm wide by 25.4 cm high. The tasks are briefly described below in the same order as they appeared in the screening battery. For more details about the tasks and scoring see Scott, Deuel, et al. (1998).
Picture pointing. Six pictures were displayed on each of four pages and children were asked to point to each picture just one time. There were two measures associated with this task. One used only the scores achieved based on the pointing sequences used by the child (quality score) and the second deducted all omissions and repetitions from the earned score (penalty score).

Picture recognition. After presenting a training memory page and test page, children were shown two sets. Each set consisted of one memory page and two recognition test pages. The examiner named each of the eight pictures on the memory page twice. On each of the two recognition test pages there were four “seen before” pictures and four new pictures. As test pages were presented, children were asked to point to the pictures that were the same as those from the memory page.

Word meaning. Children were asked, “What is an airplane?” and “What is a banana?” Examiners wrote down the children’s responses to both probes. Two measures were associated with this task.

Standard oddity. Children were asked to point to the different picture in each of nine unique arrays consisting of two identical and one different picture.

Dot matrix oddity. On each of seven pages, there were three 7.6 X 7.6 cm matrices. Each matrix was divided into nine squares. Two of the matrices had an equal number of dots placed in exactly the same position while on the different matrix, the dots were in the same positions but there was either one more or one less dot. Children were asked to point to the different design, the one not like the others.

Sequencing. On each of four pages, there was a sequence of either colored dots or animals running across the upper half of the page ending in a missing item. Children were asked to point to the one of three colored dots or animals, located on the bottom of the page, that was missing from the top.

Picture rhyming. Children were asked to point to the pair of pictures whose names sounded alike. One pair appeared above a horizontal black line and the other below it. There was one training page and two test pages. Regardless of the language used to administer the instructions, the names rhymed in English, and were labeled in English.

Unstructured semantic information. Children were first shown a grouping of cats and then a grouping of fruit. As each was presented, examiners asked the children to tell all they knew about the items displayed; cats and fruit. Examiners wrote down the children’s responses to each of the two stimuli. Two measures were associated with this task.

Structured information. Pictures of four people varying along several dimensions were displayed on a single page. Children were asked to tell how people differed from one another. Examiners wrote down the children’s responses.

Testing Procedure

All children were tested individually, for a single session, by one of six female examiners. Before starting the testing session, examiners pointed out a selection of award certificates and stickers. The children were told they could earn these by playing the game. Regardless of their performance, all children were given an award certificate and stickers. Presentation time was 15 to 20 minutes. During presentation of the test, children were periodically praised for their good performance.

Results

Task Evaluation

Children were first divided into poor or adequate achievers in reading and in math achievement. Those whose first grade percentile achievement scores were at, or below the 15th percentile were considered to be poor achievers, while those with scores above the 15th percentile were designated adequate achievers. Any child without an achievement score in the database was, necessarily, excluded.

Tasks were evaluated using frequency distributions of number of poor and adequate achievers who earned each score obtained by at least one child. Then potential cut scores were examined. All children with a score below the cut were called poor achievers while those with scores at, or above the cut were called adequate achievers. The level of accu-
racy achieved in these predictions was determined by the children’s actual achievement performance. Since we were attempting to predict which children would be poor achievers, sensitivity was defined in terms of the accuracy with which poor achievers could be identified (Yerushalmy, 1947). The cut score selected was the one that resulted in the best combination of sensitivity (percentage of all poor achievers with scores below the cut) and specificity (percentage of all adequate readers who had scores at, or above the cut).

Predicting Poor Reading Achievement

Classification accuracy. The initial evaluation was of each of the measures associated with the nine tasks. The picture recognition, standard oddity, dot matrix oddity, sequencing, picture rhyme and structured information tasks each was associated with a single dependent measure while there were two measures each for the word meaning, picture pointing and unstructured semantic information tasks. This resulted in the computation of 12 frequency distributions.

All but four children from the first grade sample had first grade reading achievement scores, so these 376 children were classified as either poor or adequate readers. Of all 12 distributions, the best combination of classification accuracy for the two groups was observed using the children’s scores on the dot oddity task where a sensitivity of 68% and a specificity of 61% was achieved. Therefore, this measure was the first selected. It was subsequently combined with the remaining 11 measures and new frequency distributions were computed for these two-task combinations. Additional tasks were added to the previous combination of tasks until no further improvement was observed in either sensitivity or specificity. The best combination available for predicting poor reading achievement was associated with a five-task combination consisting of the dot oddity, structured information (people difference), picture pointing (quality measure), picture rhyme and standard oddity tasks, using a cut score of 28. A total score summed over these five tasks yielded a sensitivity of 83% and a specificity of 60%. We accepted the slight decline in specificity in order to achieve the very high level of sensitivity. The distribution of children over this measure is shown in Table 1.

Of the entire first grade sample who had reading scores, 207 of 376 or 55% were eliminated from further consideration because they earned scores that fell at, or above the cut. This brief set of tasks reduced the number of children considered to be at risk for a reading problem to 45% of the original sample.

False positives. As can be seen in Table 1, although children’s kindergarten scores on the five-measure set led to correct identification of 83% of all the kindergarten children from this sample who would be poor achievers in first grade, there were many children with scores below the cut who were adequate readers in first grade (n = 135). These children are called false positives. They were predicted to be poor readers in first grade based on their kindergarten screening score, but they were not. Their reading scores were adequate in first grade.

We then examined the reading achievement status of these 135 children over the next two grades to see if this level of performance was stable. Might the prediction of poor achievement be incorrect for first grade but not for subsequent grades? Included in this further examination of reading achieve-

<table>
<thead>
<tr>
<th>Screening Score</th>
<th>Poor</th>
<th>Adequate</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>16–18</td>
<td>2 (n = 34)</td>
<td>6 (n = 135)</td>
</tr>
<tr>
<td>19–21</td>
<td>5</td>
<td>14</td>
</tr>
<tr>
<td>22–24</td>
<td>8</td>
<td>57</td>
</tr>
<tr>
<td>25–27</td>
<td>17</td>
<td>58</td>
</tr>
<tr>
<td>28–30</td>
<td>5</td>
<td>94</td>
</tr>
<tr>
<td>31–33</td>
<td>2 (n = 7)</td>
<td>79 (n = 200)</td>
</tr>
<tr>
<td>34–36</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td>40–41</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>41</td>
<td>335</td>
</tr>
</tbody>
</table>

Table 1

Number of Poor and Adequate Readers Who Earned Each Score on the Dot Oddity, Structured Information, Picture Pointing (Quality Score), Picture Rhyme and Standard Oddity Screening Set

Predicting Poor Achievement
ment were all children who could be classified as a poor reader or not in at least one of the two grades examined \((n = 111)\). Excluded \((n = 24)\) were those children missing reading achievement scores for both second and third grade \((n = 13)\) and those missing data for one grade \((n = 1)\) where the single recorded achievement score was greater than the 15th percentile. Such a student could not be classified as a poor achiever or not in either grade because the only data indicated adequate achievement but the missing score could have been either poor or adequate. However, if the single score was at, or below the 15th percentile, than that child was included, as a classification could be made of poor reading achievement for that grade.

Eleven of the children who were false positives in first grade were poor achievers in second grade, 16 were poor achievers in third grade and 11 were poor achievers in second and third grade. In all, 38 children, or 34% of the false positives in first grade that could be classified actually were poor achievers in subsequent grades.

If we add these 38 children to the 34 children with scores below the cut who were poor achievers in first grade (see Table 1), it can be seen that 72 of the 145 kindergarten children, or 50% of those with scores below the cut who could be classified, were poor readers in early grade school.

**True positives.** Of the 34 poor achievers in first grade who had cut scores below the cut (see Table 1), 27 could be classified as a poor achiever or not in the two subsequent grades. Of these, 18 or 67% were also poor achievers in second (8), third (5) or both second and third (5) grades.

**Predicting Poor Math Achievement**

**Classification accuracy.** For these computations we combined the two math subtests available in the data base. A child was labeled a poor math achiever if he/she earned a percentile score of 15 or less on either or both math subtests. We were able to classify 376 of the 380 children in the first grade sample. The same methods used to select measures that permitted the most accurate prediction of who would, or would not be a poor reader in first grade, were applied here to predict poor math achievement in first grade.

Children were divided into poor and adequate math achievers. Frequency distributions were computed using all 12 measures. The best combination of sensitivity (72%) and specificity (63%) was associated with the dot oddity task. The measure for this task was then combined with all remaining measures to form eleven two-task combinations, then ten three-task combinations. No further improvement was seen when combinations of four were examined. A score summed over the three-task combination of dot oddity, sequencing, and standard oddity yielded a sensitivity of 72% and a specificity of 69% using a cut score of 18. These data are shown in Table 2.

Using this subset of just three tasks, we were able to eliminate 236 children from further consideration. This constitutes 63% of the kindergarten sample. Only 37% of the kindergarten sample would be considered to be at risk for math learning difficulties.

**False positives.** As was seen for the reading data, the high level of sensitivity was associated with a large number of false positives \((n = 98)\). Once again we examined children’s performance on subsequent achievement tests to

**Table 2**

<table>
<thead>
<tr>
<th>Screening Score</th>
<th>Math Achievement Status</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Poor</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>5 (n = 42)</td>
</tr>
<tr>
<td>13</td>
<td>5</td>
</tr>
<tr>
<td>14</td>
<td>4</td>
</tr>
<tr>
<td>15</td>
<td>4</td>
</tr>
<tr>
<td>16</td>
<td>10</td>
</tr>
<tr>
<td>17</td>
<td>11</td>
</tr>
<tr>
<td>18</td>
<td>5</td>
</tr>
<tr>
<td>19</td>
<td>9 (n = 16)</td>
</tr>
<tr>
<td>20</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>58</td>
</tr>
</tbody>
</table>
see how stable this level of achievement performance was. Unlike the reading data, fewer of the false positives in first grade subsequently demonstrated poor math achievement. Of the 82 of 98 children who could be classified based on the second and third grade math achievement data, only 16% were poor math achievers in either \((n = 9)\) or both \((n = 4)\) grades. However, if we consider only those children with scores below the cut who could be classified (42 poor achievers and 13 of 82 adequate achievers), 44% were poor achievers in early grade school.

**True positives.** Of the 42 poor math achievers in first grade who had scores below the cut (see Table 2), 34 could be classified as a poor achiever or not in the two subsequent grades. Of these 34, 88% were also poor achievers in second, third or both second and third grades.

**Poor Achievement in Both Reading and Math**

Of the 41 kindergarten children who were poor readers in first grade, 21 also demonstrated poor math achievement in first grade and 20 of these 21, or 95%, earned screening scores that fell below the cut; would have been identified as children at risk for poor reading achievement.

Of the 58 children who exhibited poor math achievement in first grade, 21 were poor readers as well. Seventeen of these 21 or 81% of those who showed both poor math and poor reading had earned math screening scores that fell below the cut; were predicted to show poor math achievement in first grade.

In summary, 17 of 21 or 81% of the children who were poor achievers in both reading and math in first grade were predicted to be poor achievers in both content areas. There were 56 children who were predicted to show poor achievement in both reading and math who did not. They earned scores below the cut for both reading and math but were not poor achievers.

**Poor Achievement and Special Education Services**

We wanted to see if there was any relation between children who were classified as either EMH or SLD by third grade and their predicted first grade achievement performance. Eleven children from the kindergarten sample had an EMH or SLD label in third grade. Eight of these eleven children (73%) had been predicted to be poor achievers in both reading and math in first grade.

How did they do in achievement in first grade? Four of these eight were already in special education in first grade. There was no SAT data for two of them, one was a poor achiever in both reading and math, and one was an adequate achiever in math, but there was no SAT data on the reading subtest. All four children who were in special education in third grade but who were in regular education in first grade and earned scores below the cut in both the first grade reading and first grade math distributions, were poor achievers in both reading and math in first grade.

In summary, 8 of the 11 children from the kindergarten sample who had an SLD or EMH label by third grade had been predicted to be poor achievers in both reading and math in first grade. SAT data was available on all three subtests for only five of the eight but all five actually were poor achievers in both reading and math in first grade as predicted.

There were 13 other children from the kindergarten sample who were receiving special education services in third grade. Nine were classified as speech impaired, one as language impaired and three as emotionally disturbed. The three children classified as emotionally disturbed were all predicted to be poor achievers in both reading and math achievement. The one child classified as language impaired also was predicted to be a poor achiever in both reading and math achievement. Finally, four of the nine children classified as speech impaired were predicted to be poor achievers in both reading and math.

In effect, of the 24 children from the kindergarten sample who were receiving special education services by third grade, 67% were in that subgroup whose screening scores were below the cut in both reading and math and had, therefore, been predicted to be poor achievers in both reading and math in first grade.

**Discussion**

The major aim of this study was to evaluate the possibility that the screening performance of
kindergarten children on simple cognitive tasks might enable one to identify those children who are at risk for poor achievement in first grade. Poor achievement was defined as having an achievement score on the reading subtest or either math subtest that fell at, or below the 15th percentile.

Let us first consider the prediction of poor reading. Of all children who were poor achievers in first grade, we were able to correctly identify 83% using a screening set consisting of five tasks from a nine-task battery. This level of sensitivity is above the minimum level of 80% recommended by Meisels (cited in Lichtenstein & Ireton, 1991). When we then examined the second and third grade achievement performance of this subsample of poor readers, we saw that 67% were also poor readers in one or both grades. Clearly, this brief set of tasks was alerting us to most of the children who would demonstrate poor reading achievement in one or more of the early grade school years.

The specificity, however, was much lower. Forty percent of the adequate readers in first grade also had scores below the cut. This group of adequate readers constituted a large percentage of all children with scores below the cut. However, when we examined the second and third grade achievement performance of these false positives, it was seen that 34% were poor readers in second, third or both grades. Although their predicted designation vis a vis achievement status in first grade was incorrect, a third of these first grade false positives would evidence poor reading achievement in one or both of the next two academic years.

Although there are other kindergarten measures that have been found to relate to subsequent reading achievement, e.g., performance on certain measures from the Dynamic Indicators of Basic Early Literacy Skills (Elliot, Lee, & Tollefson, 2001) and performance on certain prereading skills (Morris, Bloodgood, & Perney, 2003), the brief time required to administer the five set task and the ease of training nonprofessionals (undergraduates) to administer them, are additional strengths of this screening set.

A similar, but not identical pattern of results was observed when we examined the accuracy with which we identified those kindergarten children who would demonstrate poor math achievement in first grade. A three-task screening set was associated with a sensitivity of 72%. That was somewhat less than we saw when predicting reading difficulties. However, the specificity, at 69%, was higher. All told, 63% of the kindergarten sample had scores at, or above the cut and could be eliminated from further consideration. Only 37% of the kindergarten sample that was successfully tracked into first grade who could be classified subsequently demonstrated poor math achievement in second, third or both grades. However, of those children who 1) had a math screening score below the cut, 2) were poor achievers in first grade and 3) could be classified as poor achievers or not in the subsequent two years, 88% also demonstrated poor math skills in second, third or both grades.

In summary, it appears to be possible to identify the majority of kindergarten children who will demonstrate poor achievement in their early grade school years. Of primary importance, the specific area of risk can also be identified, reading or math, which should be of major assistance to the first grade teachers of these children. If we only consider this aspect of the predictive characteristics of the five- and three-task screening sets, and some have argued that the most critical aspect of a screening test is its ability to identify those children who will have problems in school (e.g., Limbos & Geva, 2001; Rafter, 1997), then the subsets selected can be judged to have been successful.

Many of the false positives never demonstrated poor achievement in either reading or math. Although a high rate of overreferral is often found with developmental screening measures (Gredler, 1997; Lichtenstein & Ireton, 1984), their presence necessarily impacts how we should use the information we can deduce from kindergarten screening scores. It is obvious that one should not refer all kindergarten children with scores below the cut for a costly psychoeducational assessment and the potential anxieties created in the child and the parents that might follow such a recommendation. Rather, results of the children’s
kindergarten performance on these brief screening sets could be used to identify a sub-sample of children at risk for reading or for math achievement difficulties, a group who could be carefully monitored for the earliest signs of actual difficulties in their predicted area of risk. Then, at the earliest sign of any problems by anyone from this risk group, the combination of risk status and actual academic problems could immediately indicate the need to provide a pre-referral intervention to the target child, rather than waiting for further educational difficulties to manifest themselves. The provision, and careful monitoring of such interventions has been shown to result in academic improvement in both reading (Bursuck, Munk, Nelson, & Curran, 2002; Quiroga, Lemos-Britton, Mostafapour, Abbott, & Berninger, 2002) and math (Fuchs, Fuchs, & Kerns, 2001) even for children who are learning English as a second language (Linan-Thompson, Vaughn, Hickman-Davis, & Kouzkanani, 2003).

Monitoring, rather than placement, would incur no additional costs, nor would it lead to the improper labeling of any child. Rather, a prediction of poor achievement would simply serve as an early warning signal of potential, not certain, academic problems in a specified content area that might eventually require remedial instruction.

One alternative to monitoring all children at risk for either reading or math difficulties would be to monitor only those children whose kindergarten screening data indicate a prediction of poor achievement in both reading and math. In our data, this much smaller sample included 1) 81% of the children who actually exhibited poor achievement in both academic areas (17 of 21), those with the more severe achievement problems, 2) 73% of the children with mild cognitive impairments who were classified as either EMH or SLD by third grade, and 3) 67% of all those children from the kindergarten sample who were receiving some type of special education service in third grade. In addition, the number of false positives was much smaller \((n = 56)\), requiring the monitoring of only 23% of the kindergarten sample.

Not included in this smaller sample, however, would be 71% of the children who will evidence only poor math achievement in first grade and 51% of the children who will only be poor readers in first grade.

Another alternative would be to use these sets as initial screens to reduce the total number of kindergarten children who would need to be screened with more lengthy and perhaps curriculum tied screening instruments.

These initial data were developed on a sample with their own specific characteristics. Much further work is required, for example work on cross validation and estimates of reliability. However, they do indicate the possibility that one might be able to use brief cognitive screening sets to identify a group of children in kindergarten who are at risk for a specific area of achievement difficulty in the early grade school years and perhaps also those at risk for even more involved educational problems who will ultimately require special education services. Monitoring the actual performance of this limited subset of children in order to detect any early signs of learning difficulties would be an inexpensive method of attacking a major educational problem.

References


Funk, S. G., Sturner, R. A., & Green, J. A. (1986). Preschool prediction of early school performance: Relationship of McCarthy Scales of Children’s Abilities prior to school entry to achieve-


Swanson, H. L. (1994). Short-term memory and working memory: Do both contribute to our understanding of academic achievement in children.


Received: 2 June 2004
Initial Acceptance: 1 August 2004
Final Acceptance: 11 January 2005