



STUDY TITLE:
Screening With the TPRI/AMIRA Screeners

Author: Dr. David Francis, University of Houston

Date: 2016



Research On Amira Over Time

It is well-established that the key to intervening with children at-risk for dyslexia is early intervention (Fletcher et al., 2018). When identified early- in kindergarten, Grade 1, and Grade 2, the number of children who do not respond to explicit core reading instruction, often with supplemental reading intervention, is low; many studies reporting reductions of risk from 20% to under 5% of children depending on the quality and intensity of the interventions (e.g., Mathes et al., 2005). In contrast, when intervention is delayed until the reading problem is fully manifested, intervention is much less efficacious. Studies report that beginning in third grade, more time in intervention is required to accelerate gains compared to grades 1-2 (Lovett et al., 2017). The reason is that the neural systems that must emerge in order to support proficient reading require considerable exposure to print. If a child is delayed in their access to print, as is the case for many children with dyslexia, these systems do not receive the experience needed to create the expertise for rapid processing of print (Seidenberg, 2017). It is difficult to provide sufficient experience if a child goes through 1-3 years of school with limited ability to read (Torgesen et al., 2001).

Because of the importance of early intervention, 37 states now mandate screening for dyslexia in kindergarten, Grade 1, and sometimes Grades 2-3. But these laws commonly confuse screening and diagnosis. Screening is rapid triage of entire classrooms to identify risk. Diagnosis is a more extensive assessment that can be costly and time consuming for teachers and other school personnel. A screen minimizes demands on teacher time. A diagnostic assessment should be done with children who are at risk on the screen. The technology needed for early screening began to emerge almost 40 years ago with the development of assessments designed to predict which children will develop dyslexia based on kindergarten assessment (Benton & Pearl, 1978). Since that time, scientific knowledge has accumulated in terms of what needs to be assessed, psychometric approaches to test construction and decision theory, and how to make screening efficient in schools, which has led to considerable reduction in the time required for screening. In 1997, Texas passed a law mandating early screening for reading problems, including dyslexia, in Kindergarten and Grades 1-3. The Texas Education Agency contracted with the Center for Academic and Reading Skills (CARS) and the Texas Institute of Measurement, Evaluation, and Statistics (TIMES) to develop a screening instrument and an inventory to meet the requirements of this law. The result was the Texas Primary Reading Inventory (TPRI/AMIRA), which provided a 3-5" screen to identify children at risk for reading problems and a 30' inventory to determine what reading concepts needed to be taught. The TPRI/AMIRA underwent a significant period of development, with psychometric studies in 1997-2000 and an updating in 2010. It remains in use and the screening portion is largely unchanged since its initial development.

The screening component of the TPRI/AMIRA was unique and built on a long history of designing screening instruments dating back to the Florida Longitudinal Project (1970-1978; Satz & Fletcher, 1978). Drs. Fletcher and Francis developed a kindergarten readiness assessment for the Houston Independent School District in 1980-1985. The NICHD-funded Early Assessment of Reading Skills (EARS) project by Dr. Francis, Dr. Foorman, and Dr. Fletcher lead to the development of the TPRI/AMIRA.

From these studies, we understand that the most important component of a screening instrument is predictive validity. Children change rapidly during their first few years of schooling. Screening that is concurrent, i.e., addresses only the child's status at the time of the assessment, does not take into account these changes and cannot be validated as predictive of the child's status in subsequent grade. EARS was a longitudinal assessment of early reading skill in over 900 children unselected for reading problems. It had 4 cognitive assessment time points in Kindergarten, grade 1, and grade 2, with reading outcomes at Grades 1 and 2. We were able to use EARS to build and initially validate the TPRI/AMIRA screening instruments.



Research On Amira Over Time

We then did additional evaluations of reliability and validity in different large samples in 1998-2000 and 2010. In all these studies, the screening instruments were reliable and had strong predictive validity.

In addition to the capacity for predictive validity, screening instruments must not be confused with diagnostic instruments. The purpose of a screen is to identify children who may be at-risk as quickly and efficiently as possible. The screen should be followed by an inventory, a diagnostic test, or a progress monitoring system to delineate what to teach and whether the child's learning is accelerating to reduce their risk. All screens yield errors. Some children will be identified as "at-risk" who will not go on to develop dyslexia (false positive error). Other children will be missed, i.e., identified as not "at-risk" when they in fact develop reading problems later in development. These errors depend on where thresholds are placed and are related: as false positives go up, false negatives go down, and vice versa. Therefore, it is important to determine what types of errors are less desirable and place the threshold to minimize these types of errors.

In developing the TPRI/AMIRA screening instruments, we determined that we wanted the screens to take less than 5' for the teacher to administer. So we worked with the EARS data and identified tasks that were most predictive of subsequent reading problems: alphabetic knowledge and phonological awareness in Kindergarten, phonological awareness and word reading in Grade 1, and word reading in grade 2 (Schatschneider et al., 2004). We then used item response theory to identify the items that were most predictive and that clustered around a threshold that predicted risk status. We also examined whether items performed differently by gender and ethnicity (differential item functioning) and further refined the items. We determined that since we were developing a screening instrument, false negative errors were much less desirable than false positive errors because the consequences of missing an "at-risk" child and delaying access to intervention were more serious than a false positive error, in which the child's development of reading skills would show the error as development was monitored or through performance on a diagnostic test. The value of this approach is that a 3-5" screen can be done quickly with the entire class and will reduce the number of children who need monitoring or diagnostic testing. The screen is highly accurate for identification of children who are not at-risk for dyslexia. They do not require further assessment, but the screen should be periodically administered to the entire class to ensure that children who are at-risk are not missed. False positive rates are higher in kindergarten, but decline considerably through grade 2.

These efforts were successful. There are five different screens designed for administration in the middle and end of kindergarten predicting to end of grade 1; beginning and end of Grade 1 predicting to end of grades 1 and 2, and beginning of Grade 2 predicting to end of Grade 2. There is no beginning of kindergarten screen because children require time to acclimate to school. All of the screens were developed to minimize the errors that result from not identifying children who need further assistance. Thus, the screens provide brief assessments (3-5 minutes) that permit identification of children who are not likely at risk for the development of reading difficulties. Children who meet the criteria on the screen do not require further assessment. Teachers should also use their judgment and experience with the child to further evaluate the accuracy of the screening information.

Each of the five screens developed for the TPRI/AMIRA in 1997-2000 fail to identify less than 10% of the children who end up not reading at expected levels by the end of Grades 1 and 2 (false negative errors). Even when the goal is to identify children who are not at risk, errors involving the over-identification of these children are inevitable (false positive errors). However, errors of this type are viewed as less serious than failing to identify children who are at risk. Moreover, false positive rates were uniformly below 45% for Kindergarten and Grade 1, and dropped dramatically by the second grade (15%).

In the 2010 study, the screens were re-evaluated. A Grade 3 screen (word reading) was added and the



Research On Amira Over Time

initial kindergarten assessment was moved to earlier in the year. All components had reliability coefficients ranging from .88-.93. There was little evidence of item bias. Predictive validity was examined from screens at the beginning of the year to the end of the year. Measures were collected in both the fall and spring at every grade level. Outcome measures were administered at the end of the year and students were evaluated against whether they fell above or below a set threshold on the outcome. Balancing correct identification and minimizing false-negative identifications is the primary importance of the TPRI/AMIRA screens. Across the three grades and six screen forms (K-MOY, K-EOY, G1-BOY, G1-EOY, G2-BOY, G3-BOY), we evaluated the revised screens with data from 4581 student outcomes. While the TPRI/AMIRA screens would have correctly identified over 70% of those students, it is more instructive to consider that the TPRI/AMIRA screens would have failed to identify only 46 out of 4506 (~1%) students through the use of short assessments that take less than 3 to 5 minutes per student.

The TPRI/AMIRA screening instruments are unique. Unlike other proposed screening devices, the TPRI/AMIRA has predictive validity and is designed to identify risk level and forecast subsequent development of reading problems, including dyslexia. It has strong reliability and validity, requires 3-5', and very reliably indicates students who are not at risk. These students do not need further assessment depending on the teacher's judgement. If a child is identified as "at risk," further assessment and monitoring is warranted. Early detection of reading problems is the key to prevention dyslexia. By rapid screening and triaging of children according to risk status before reading problems become fully manifested, the TPRI/AMIRA screening instruments are pivotal to early intervention and the need to comply with state laws for early screening of dyslexia. There are many options for children who test positive, including inventories, diagnostic tests, and progress monitoring.



Research On Amira Over Time

REFERENCES

- A. L. Benton & D. Pearl (Eds.), *Dyslexia* (pp. 123–137). New York: Oxford University Press.
- Fletcher, J.M., Lyon, G.R., Fuchs, L.S., & Barnes, M. (2018). *Learning disabilities: From identification to intervention*. New York: Guilford Press.
- Lovett, M.W., Frijters, J.C., Wolf, M.A., Steinbach, K.A., Sevcik, R.A., & Morris, R.D. (2017). Early intervention for children at risk for reading disabilities: The impact of grade at intervention and individual differences on intervention outcomes. *Journal of Educational Psychology*.
- Mathes, P. G., Denton, C. A., Fletcher, J. M., Anthony, J. L., Francis, D. J., & Schatschneider, C. (2005). An evaluation of two reading interventions derived from diverse models. *Reading Research Quarterly*, 40, 148–183.
- Satz, P., Taylor, H.G., Friel, J., & Fletcher, J.M. (1978). Some developmental and predictive precursors of reading disability. In A.L. Benton & D. Pearl (Eds.), *Dyslexia: An appraisal of current knowledge* (pp. 457-501). New York: Oxford U. Press.
- Schatschneider, C., Fletcher, J. M., Francis, D. J., Carlson, C. D., & Foorman, B. R. (2004). Kindergarten prediction of reading skills: A longitudinal comparative analysis. *Journal of Educational Psychology*, 96, 265–282.
- Torgesen, J. K., Alexander, A. W., Wagner, R. K., Rashotte, C. A., Voeller, K. K. S., & Conway, T. (2001). Intensive remedial instruction for children with severe reading disabilities: Immediate and long- term outcomes from two instructional approaches. *Journal of Learning Disabilities*, 34, 33–58.