Spend any time observing literacy instruction in a preschool or primary grade classroom and you cannot help but be struck by the range of individual differences in learners. In a single first grade classroom, you might find one child who has mastered letter/sound pairings and recognizes many sight words. This child easily reads aloud a book meant for older children, confidently answering questions about the storyline. A second child may have strong language skills and love listening to stories, but struggles to name letters and pair them with their sounds. A third child may be encountering English for the first time, speaking a different language at home. This child may have a large vocabulary in another language but not know the meaning of everyday English spoken words such as “today” or “summer.”

Educational research often validates reading practices for particular types of learners, such as children reading on grade level, children at risk for learning disabilities, or children who are English Language Learners (ELLs)(1-3). But what if all these children are in the same class? Knowing to whom research findings apply and then figuring out how to integrate research-based practices into a busy classroom where children have very different needs can be difficult. The dilemma of which research-based practices to implement, and for whom, is one facet of the implementation gap that often prevents research-based practices from improving outcomes for children when they are implemented in real world classrooms(4).

Individual differences among students are not the only aspect of classroom life that contributes to the research implementation gap. It can be hard to fit research-based practices to a school’s specific circumstances. A research-based practice that fits easily in one first grade core curriculum may not work well in a second curriculum that introduces skills in a different sequence. A research-based practice that works well in a classroom with a teaching assistant may not be possible in classrooms with a single teacher(5).

In sum, improving literacy outcomes for all children is not just a matter of finding out what works, but understanding what research-based instruction works for which
child and then providing guidance in **how those practices can be implemented in diverse classroom contexts**. In the next sections, we summarize the innovative research of Carol Connor and her colleagues. Across two decades, Dr. Connor first used observational research to investigate how the individual differences that children bring into the classroom can result in very different outcomes for children when receiving the same instruction. Connor then used this knowledge to develop a flexible tool, **Assessment to Instruction, or A2i**, that helps schools deliver research-based literacy instruction matched to the unique needs of individual learners to increase reading achievement for all students.

The first part of this summary describes Dr. Connor and colleagues' early work observing reading instruction and documenting that the same instruction can produce very different outcomes for different children. The second part describes the development of the A2i intervention and the results from the multiple randomized controlled trials (RCTs) demonstrating that A2i improves literacy outcomes for children. We end with a summary of current efforts to bring A2i to scale so that it can be widely used in schools and classrooms. This work has been driven by one overarching goal: every child reading at, or above, grade level by the end of third grade.

**PART ONE: LOOKING IN CLASSROOMS TO UNDERSTAND WHAT WORKS**

Starting in the early 2000s, Dr. Carol Connor and her colleagues spent thousands of hours over five years observing literacy instruction in preschool through third grade classrooms. The observations were recorded in minute-by-minute observer narratives and later on videotapes (6-9). Everything that went on during literacy instruction was recorded. The beauty of starting with observations that simply documented everything happening in the classroom was that the observations could be taken back to the lab and coded many different ways. Final decisions about what aspects of instruction were most important in children's learning could be made by statistically examining how different dimensions of instruction predicted children's literacy achievement.

Three additional features of these observational studies had important implications for understanding how reading instruction influences children's reading growth.

- In most studies, Connor and colleagues observed instruction in the **fall, winter, and spring** of school years. This strategy allowed them to examine whether instruction changed over the course of the year. Children gain skills and become more independent over the school year and these changes can be expected to influence children's response to different types of instruction.

- Connor and colleagues administered measures of children's vocabulary, decoding skills, and in later grades, reading comprehension, at the **beginning of the year and at the end of the year**. Assessing children at least twice allows measurement of student growth. Without knowing children's entry level skills,
we might attribute high achievement at year end to good instruction even if the
student entered above grade level and had gained little across the year. We
might also miss good instruction that had lifted a student from two years below
grade level to within half a year of grade level because the child was still
somewhat behind his or her peers at the end of the year.

• Finally, Connor and colleagues looked at the impact of instruction on groups
of students with different entry level skills. Why is this important? Because
only looking at growth for students on average means we cannot see how
diverse students in the classroom respond to the instruction they get. The three
students described in the opening paragraph of this report have very different
instructional needs. But their different responses to the same instruction would
be obscured without considering how the instruction interacted with their
entry-level skills. Unless researchers consider the impact of instruction on
individuals with different profiles of skills, they will only find a “one size fits all”
model of good instruction. Those practices may produce good outcomes on
average but are likely to fail many learners.

Roots in the Science of Reading and Instructional Research
As Dr. Connor and colleagues began to analyze the classroom observations, they
were guided by earlier research and theory on two topics: the component skills of
reading, often referred to in the mainstream media as “the Science of Reading,” and
research examining the elements of effective instruction. The Simple View of Reading
proposed by Gough and Tumner(10) represents a useful way to summarize many of
the findings in the science of reading and also think about the impact of students’
entry level skills on their reading progress. This model posits that reading
comprehension is the product of students’ decoding skills and their linguistic
comprehension (i.e., decoding X linguistic comprehension = reading
comprehension). Students acquire decoding skills so they can translate written text
to oral language, and once text is decoded, their linguistic skills allow them to
understand what they have read. Weaknesses in either set of skills will result in poor
reading comprehension, and students vary in the ease with which they acquire either
set of skills. This model, along this the critical components of instruction (phonemic
awareness, phonics, fluency, vocabulary, and comprehension) delineated by the
National Reading Panel in 2000 served as the jumping-off point for Dr. Connor’s work.

The Simple View of Reading is a useful way of thinking about role the five
components of instruction delineated by the National Reading Panel(11) play in skilled
reading. Students must understand that spoken words can be broken down into
sounds (phonemic awareness) that are represented by letters (phonics). Beginning
readers then use this knowledge to rapidly move back and forth (fluency) between
the spoken and written forms of sounds and words. That is, readers need to master
the code that links written and spoken forms of language. For example, beginning
readers need to be able to say “cat” when seeing the letters “c,” “a,” and “t,” and also
to write the word “cat” when hearing it spoken aloud. Although decoding skills are
important, linguistic skills are also critically important to reading—children must have
or acquire the background knowledge and spoken vocabulary knowledge to understand the meaning of words and passages once they had been decoded (vocabulary and comprehension). If you’ve ever read aloud a passage in a foreign language and had no idea what you’ve read, you can appreciate that pronouncing words correctly and understanding their meaning are very different things.

In addition to research on component skills of reading, Connor and colleagues used earlier research on effective instruction to guide their work. This body of research had indicated that time was a very important element of instruction and could be helpful in understanding the effects of instruction. For instance, an early study that tracked how much time teachers spent on different topics revealed that in the upper elementary grades, teachers spent very little time teaching reading comprehension. Tracking how many minutes of instruction individual children got in the component skills of reading versus simply looking at whether or not key components of instruction were covered was an important research innovation. Tracking minutes in key components of instruction allowed close examination of the impact of instruction on individual students.

Other work in instructional research indicated that different teacher-student groupings have different effects on student learning. When children worked one-on-one or in small groups with the teacher guiding them, they typically attended better, content was more likely to be targeted to their needs, and new skills were more easily acquired. But there was also research showing that as children moved beyond initial skill acquisition, it benefited them to interact among themselves and to practice the skills on their own.

After applying different coding schemes to their classroom observations, Connor and colleagues ultimately settled on a set of four variables that formed a two by two matrix of important types of instruction. Although their final classification scheme was simpler than some of the coding systems that they had initially tried, it was quite useful. Consistent with the Simple View of Reading, they divided the ‘content of instruction’ into two types, meaning-focused and code-focused. Consistent with instructional research, they chose to focus on two instructional groupings, distinguished by who managed children’s attention in learning—the child or the teacher. By identifying both the content of instruction and who was managing the students’ learning, reading instruction could be placed into one of four quadrants. Table 1 provides examples of each of the four types of instruction. Although this was a very simple taxonomy, much of children’s reading growth could be predicted simply by knowing how much of each type of instruction children had received across a school year. The final coding system for classroom observations tracked the amount of time in each of these four areas.
Table 1. Examples of the four types of reading instruction in Connor’s final observation system

<table>
<thead>
<tr>
<th>CONTENT OF INSTRUCTION</th>
<th>DIMENSION OF INSTRUCTION</th>
<th>Teacher Managed (TM)</th>
<th>Child Managed (CM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Code-focused (CF)</td>
<td></td>
<td>The teacher is working with a small group of children on how to decode compound words such as “cowboy” and “baseball.” She says: “What word do you have when you take the ‘boy’ out of ‘cowboy’?”</td>
<td>Children are working together to sound out and then write words that have the rime “-ake.” They have written the words “bake” and “cake” on the white board.</td>
</tr>
<tr>
<td>Meaning-focused (MF)</td>
<td></td>
<td>The teacher is discussing the story “Stone Soup” with the class. She starts by asking the children: “What is the main idea of the story and what are the supporting details?” She then tells them to: “Think, pair, share” so the children turn to their partner to discuss the main idea and supporting details.</td>
<td>Children are silently reading a book of their choice at their desks. Other children are writing in their journals.</td>
</tr>
</tbody>
</table>

Drawing on Research and Theory About Individual Differences

As stated in the introduction, it is obvious that children, even in a single classroom, can vary on many characteristics important to learning. In examining learners in their early observational research, Connor and colleagues looked at a number of child characteristics. Consistent with the Simple View of Reading, they ultimately found that two child characteristics were quite important in determining how a child read at the beginning of the year and how much progress the child would make by the end of the year. Those two characteristics were the children’s oral language skills (e.g., knowledge of the meaning of spoken words often measured with a vocabulary measure) and their facility with the code system that links spoken and written words. Of course, other child characteristics were important, too, but a child’s level on these two skill sets were the most useful in predicting children’s progress in early reading. They are also two characteristics that are malleable—that is, influenced by instruction.

1 This table is adapted from Connor C. (2014). Individualizing teaching in beginning reading. Better: Evidence-based Education. 6(3): 4-7.
Putting It All Together: Finding What Matters in Reading Instruction

The development of the observational coding system and selection of two child characteristics important in predicting reading outcomes were the means to accomplish the most important aspect of Connor’s observational work. She was then able to examine how children’s entering skill levels interacted with the patterns of instruction they received. In each observational study, Connor and colleagues found child X instruction interactions (read as ‘child by instruction interactions, meaning a statistical interaction between two variables). That is, children’s reading progress was best predicted by looking at how the instruction they received fit with the unique constellation of skills they brought to the classroom. They found child X instruction interactions in each grade, although the nature of those interactions changed with grade level.

In this section, we provide two examples of child X instruction interactions. These examples are taken directly from Connor and colleagues’ research. They illustrate why it is critical to abandon the assumption, implicit in many theories of instruction, that that there is one type of effective reading instruction for all children. When the characteristics of individual learners are ignored, students can be expected to make less academic progress than if the instruction was matched to their entry level skills. Children who are fortunate enough to be placed in a classroom that fits their unique set of needs will make more gains than children who are placed in a classroom where their entry level characteristics are a poor fit with the instruction they receive. Child X instruction interactions affect literacy outcomes for children at all levels of reading skill, not just low achievers or children at risk.

Both examples in this section are from Connor’s study of first grade classrooms(6). They illustrate the importance of considering children's skill development when designing reading instruction, and how reading gains can vary when there is a good or poor match of instructional patterns with children’s skills.

Child X Instruction Interaction: Example 1

This first example, depicted in Figure 1, illustrates how the same amount of child-managed instruction can be associated with very different outcomes for first graders who enter the grade with different levels of vocabulary knowledge. The x-axis is the average amount of child-managed instruction observed per day in different classrooms (e.g., silent reading, journaling, students working in pairs, workbooks). As is evident in the figure, classrooms varied widely in how much time children spent working on their own (16-37 minutes, see numbers on the x-axis). The y-axis is children’s decoding score in the spring—their word identification skills as measured on a nationally standardized test. For students who entered first grade with a high level of vocabulary knowledge (scoring better than 75% of their peers), more time spent in child-managed instruction was associated with better reading outcomes. This finding is indicated by the steep incline in the solid black line depicting how reading outcome varied at different amounts of minutes spent in child-managed instruction. Conversely, for students who entered with low vocabulary skills (scoring
better than only 25% of their peers), more time spent in child-managed instruction was associated with lower word reading skills, as indicated by the steep decline in the dotted line linking reading outcomes and minutes of child-managed instruction for this group. Children who were in the average range (50th percentile) in vocabulary skills had similar word reading skills regardless of the minutes of child-managed instruction. Note that only at about 22 minutes of child-managed instruction per day did the three lines showing the relationship between minutes in child-managed instruction and decoding scores for the three groups of students with different entry decoding skills intersect, indicating similar word reading outcomes for all groups of students.

**Figure 1. Child X Instruction Type interaction effect for spring decoding raw scores. Child fall vocabulary raw score (25th, 50th, and 75th percentiles of the sample) by CMI amount**

![Graph showing the interaction effect between child fall vocabulary and CMI amount on spring decoding scores]

**Child X Instruction Interaction: Example 2**

This second example, shown in Figure 2, illustrates the extent to which the “fit” of an instructional pattern with a child’s needs can be associated with better or poorer reading outcomes. In this example, gains in decoding skills are contrasted but only for the subgroup of first graders who had low fall vocabulary scores and low decoding skills (25th percentile for each skill). Although the entire group started at the same place in the beginning of the year (scoring about half a year below grade level), where children ended the year varied systematically with the pattern of instruction they received during the year. Children who got smaller amounts of teacher-managed explicit instruction (e.g., instruction in word segmentation and letter sounds)
and greater amounts of child-managed implicit instruction (e.g., reading or writing on their own or in groups) made very modest gains and ended the year further behind their peers. The decoding skill gain for children who got this ineffective pattern is depicted in Figure 2 with the blue dotted line. Children who received substantial amounts of teacher-managed explicit instruction, paired with smaller amounts of child-managed implicit instruction that increased across the year, made large gains in decoding skills. By the end of the year, the students who got this pattern now scored an average grade equivalent of 2.5 in decoding. The decoding skill gain for students who got this more effective pattern is depicted in Figure 2 by the green line.

It is important to note that these particular patterns of instruction were ineffective and effective for this group of students only. For example, students who entered first grade with higher levels of decoding and language skills made more growth in decoding skills if they had greater amounts of child-managed implicit instruction. Thus, patterns of types of reading instruction across the year were only effective and ineffective relative to children’s initial decoding and language skills.

**Figure 2.** Children who began the school year with low fall vocabulary raw scores (25th percentile standardized norms) and low fall decoding raw scores (25th percentile standardized norms) but got different amounts of teacher-managed explicit instruction and child-managed implicit instruction, showed very different gains in their decoding skills from the beginning to end of first grade.
PART TWO: MOVING BEYOND OBSERVATIONAL RESEARCH TO FIELD TRIALS

One shortcoming of observational research is that the findings are only correlational\(^{17}\). In the case of Connor and colleagues’ research, less and more effective patterns were determined retrospectively, after the observations were collected and the coding system refined. Much stronger evidence for the importance of child X instruction interactions for improving reading outcomes requires an experimental design where classrooms are randomly assigned to receive typical instruction or individualized instruction. To demonstrate their findings were robust, Connor and colleagues needed to show the results they had found for different learner groups in the observational research could improve child achievement outcomes when they worked with teachers to make their instruction more responsive to individual learners.

Creating the Grade Level Algorithms

The first step in this process was to create and test specific algorithms that would translate the results of the observational research into actionable information for teachers\(^{18}\). In keeping with the importance of amount of time spent in different activities to learning, the actionable information provided for teachers would take the form of recommended minutes per day of instruction for a child in the four types of reading instruction depending on the child’s entry level skills.

Figure 3 shows an example of what these algorithms look like for students scoring at different grade levels in reading at the beginning of first grade. Please note the horizontal axis depicts a student’s reading level at the beginning of first grade, not their actual grade level (which would be first grade, first month for all students). A first-grade student entering the school year reading at a second-grade level would be most likely to benefit from slightly less than 10 minutes of adult-managed code focused instruction, 10 minutes of child-managed code focused instruction, 15 minutes of child-managed meaning-focused instruction, and 20 minutes of adult-managed, meaning focused instruction. Note that a student who was on grade level, scoring at beginning first grade level would benefit from a very different mix of the four types of instruction.

Another way of looking at this figure is to consider a first grader who starts the year on grade level for word reading needs. He or she needs only 10 minutes per day of adult-managed code-focused skill instruction (the red line). But if they are behind by half a school year, the child would need 20 minutes a day of this type of instruction. If the child is behind by a full year, they would need 30 minutes per day. These recommendations about the number of minutes of different types of instruction change over the course of the year as a student’s literacy skills grow and he or she matures. That is why the lines in the figure curve. This chart illustrates the algorithms for children entering first grade. The algorithms change at each grade because the observational results were different at each grade.
Figure 3. Recommended amounts of the four types of instruction for first graders with reading skills varying from kindergarten to third grade at the beginning of first grade.

Helping Teachers Use the Algorithms
As noted in the introduction, an implementation gap can impede translating research into classroom results. When improving achievement depends on using assessment scores to modify day-to-day instruction, teachers can struggle knowing what materials match up with the assessment results for a child. Some assessment tools link assessment results to instruction by creating their own instructional materials. But linking assessments only to new materials presents problems when teachers have to abandon practices they know well and have used with their students. To avoid this pitfall, Dr. Connor took an unconventional step: indexing already developed instructional materials to the recommend minutes of instruction for each of the types of reading instruction in her categorization scheme. That way, teachers could follow the recommendations suggested by the A2i algorithms, but make use of the materials already present in their classrooms. This strategy decreased the teacher time and district money spent on new materials. Dr. Connor also indexed high quality instructional materials available free to districts, such as the instructional materials downloadable from the Florida Center for Reading Research.

Another aspect of making the algorithms usable for teachers was to create a platform where children’s assessment results could be reviewed online. The platform also included tools for planning instruction based on each child’s assessment results and recommended minutes of instruction. For example, the platform grouped students with similar needs into small groups for more individualized instruction and suggested lessons from the classroom’s core reading curricula or from additional instructional materials.

---

2 The recommendations graphed in Figure 3 represent a grade 1 student with decoding skills ranging from kindergarten to grade 3. The impact of vocabulary level is held constant across this range. The data graphed represent a student entering grade 1 with vocabulary skills typical of a ‘on grade level’ English-speaking student.

3 Today the algorithm is indexed against over 100,000 instructional activities, including many different literacy curricula.
online resources targeted toward the small group’s specific needs. The web-based planning tools allowed the teacher to adjust instructional groups and minutes of instruction needed in the four types of instruction as children progressed. Although the A2i platform had the advantage of allowing teachers to use familiar curricular materials, it still required teachers to learn new software, implement new grouping strategies, and add lessons where their existing curriculum had gaps. To support teachers as they learned to use the software and modified their instruction, a professional development and ongoing coaching model also was incorporated into A2i. Collectively, all of these aspects of A2i are now called the A2i Professional Support System.

Validating A2i With Rigorous School-based Research
With the algorithms, a web-based platform, and professional development in place, the first randomized controlled field trial of the A2i Professional Support System was conducted. The first study examined the system’s use in first grade only because findings in other grades were still accumulating from the observational studies. The initial field trial results, published in the journal *Science* in 2007(18), were positive. In an economically and ethnically diverse school district in Florida, first grade teachers using A2i to individualize instruction showed stronger reading growth in their students than teachers who did not use A2i. A second study published in 2011, again using first graders but conducted in second school district, also showed that schools using A2i had better reading outcomes(29). As more experimental field trials of A2i were conducted(20-23), the web-based platform continued to be strengthened and refined. It was expanded for use in other grades (kindergarten, grades 2 and 3) and modifications to the tools and platform were made based on teachers’ use and feedback. Later, as a result of more research, online assessments of decoding and language skills were added to the system so that the algorithms could be used without purchasing additional tests and teachers could streamline time spent assessing students and scoring tests.

A particularly important study of A2i’s effects was published in 2013(24). The study followed students who been provided A2i in 0, 1, 2 or 3 years while they progressed from first to third grade. The study showed that reading gains continued to accumulate for children when A2i was provided in multiple grades. Students who had been in A2i classrooms for three years in a row showed average reading achievement results well above grade level in third grade (grade equivalent of 5.0 on a nationally standardized reading test).

Including the first experimental field test of A2i published in 2007, Dr. Connor and her colleagues have now published seven articles summarizing the results of randomized controlled trials of A2i. These rigorous studies demonstrate the efficacy of their strategy for using assessment results to guide instruction, and the importance of moving beyond a one size fits all model of reading instruction. Three of the studies have been reviewed by the *What Works Clearinghouse* and have met their standards without reservations(21, 23, 24). The references for the RCTs of A2i are provided in Appendix A, along with a complete list of publications by Dr. Connor.
**English Language Learners**

Non-native speakers of English often learn how to decode words without being able to fully comprehend what they are reading, due in large part to limited vocabulary skills. However, the initial RCTs of A2i that were conducted had relatively few students with limited English proficiency (a group also known as English Language Learners or ELLs). Later, Dr. Connor did additional research with students who were ELL to make adjustments in the A2i algorithm to account for their unique needs. Figure 3 below shows the results of a yet-to-be published study of the use of the adjusted A2i algorithm with ELL kindergarten and first grade students in Arizona.

Depending on their level of proficiency with English, students who were ELL were in one of two types of classrooms: mainstreamed with students whose first language was English or in Structured English Immersion classrooms where there was a focus on explicitly teaching English as well as the state’s English Language Arts curriculum. In half of both types of classrooms, teachers began using A2i. In the other half, teachers used their usual teaching methods. In the fall, there were no differences between students assigned to A2i or business as usual classrooms. Figure 4 shows the spring Word Recognition scores on a nationally normed test for students who did and did not receive A2i instruction in the two types of classrooms. The blue and green bars on the left show the decoding skills in classrooms in which teachers used A2i. The blue and green bars on the right show the results in control classrooms where it was business as usual. The green bars compare outcomes for students with and without A2i who were in mainstream classrooms. The blue bars show the comparison for A2i and non-A2i classrooms in Structured English immersion settings. A2i-informed instruction was more effective for children in both types of classrooms. The differences in outcomes for the A2i and non-A2i classrooms were statistically significant.

**Figure 4. Effect of use of A2i in classroom instruction for English language learners**

![Figure 4](image-url)
PART THREE: BRINGING A2i TO SCALE

As the results outlined throughout this document demonstrate, A2i’s foundation clearly stems from research, but is now primarily used in non-research settings. This jump, from research to practice, was made possible through funding by the U.S. Department of Education in the form of an Education Innovation and Research (EIR) Expansion grant(26). The EIR Expansion grants are awarded to companies working to bring rigorously tested research-based programs into the mainstream education marketplace. The A2i Professional Support System is currently deployed in over 100 schools across the United States, actively bridging the gap between research and practice. The implementation of A2i is supported by Learning Ovations, a company whose mission is to get all students reading at, or above, grade-level by the end of third grade.

To ensure this mission becomes a reality and the implementation gap between research and practice successfully bridged, districts using the A2i software platform also are provided access to Literacy Outcomes Specialists, coaches who help teachers move from assessment results and algorithms to actionable next steps in making their instruction more responsive to individual students. Although the path continues to be an exception, not the norm, A2i serves an example of just how the research and practice gap can be addressed to the benefit of districts, teachers, and most importantly, students.
REFERENCES


25. C. M. Connor et al., Scaling up Assessment-to-Instruction (A2I) technology to support individualized instruction with monolingual and dual language learners: Feasibility and promise of efficacy. (submitted).

Appendix A

The following studies represent the main research publications related to the impact of Dr. Connor’s research as it relates to individualizing student instruction and the A2i Professional Support System.


See below for a more complete list of publications by Dr. Carol Connor:


doi:http://dx.doi.org/10.1080/10573569.2015.1021060


