This is a chapter excerpt from Guilford Publications.
Differentiated Literacy Instruction in Grades 4 and 5: Strategies and Resources, Second Edition.
Sharon Walpole, Michael C. McKenna, Zoi A. Philippakos, and John Z. Strong. Copyright © 2020.
Purchase this book now: www.guilford.com/p/walpole4

CHAPTER 4

Choosing and Using Reading Assessments

Copyright © 2020 The Guilford Press

Our understanding of how to use reading assessments quickly to guide differentiation is the focus of this chapter. We address writing assessments in Chapter 9 as we describe Tier 1 writing instruction. We suspect that many classroom teachers view the idea of reading assessment as a complex and slightly mysterious process—one that involves giving an assortment of tests and applying daunting inferential strategies to arrive at diagnostic conclusions. Although there is some truth to this perception with regard to a small number of challenging cases, we wish to allay any fears you may be harboring about the assessment required to make small-group instruction successful in an upper elementary classroom. This perception may also stem from the fact that many teachers are forced to give assessments that are not actually used to make decisions. Others must give assessments that actually produce conflicting information. The fact is, only a few informal assessments are needed to group students for differentiated instruction and to gauge their growth.

TYPES AND PURPOSES OF ASSESSMENTS

Assessment serves a critical role in differentiated reading instruction. It guides the process first by grouping students with similar needs, then by helping to plan instruction, and finally by gauging the extent of student learning. A distinction is sometimes made between assessment *for* instruction and assessment *of* instruction. When we assess for instruction, we use the information we gather to target our teaching toward student needs. Assessment of instruction is conducted to determine whether our efforts have been successful. In our approach to differentiated instruction, both kinds of assessment are important. The fact that assessment has multiple purposes has led to more than one type of assessment, and we begin with an overview of the four basic types.

Screening measures provide an indication of achievement in a particular area. Screenings are common for word recognition, fluency, and comprehension. They are sometimes administered individually and sometimes to an entire classroom at once. These measures are limited in what they can tell us. Identifying a problem area is a good first step, but it does not suggest specific actions we might take to address the problem. To accomplish that aim, we administer *diagnostic measures*. These are follow-up tests that break down the area into teachable skills and strategies. For example, if a screening measure of word recognition indicated that a student was performing below expectations in general, a follow-up inventory of specific decoding skills would provide the information needed to identify and address the specific deficit. Screening and diagnostic measures work in tandem to provide the information teachers require in order to meet their students' needs. Our approach to differentiation in small groups for upper elementary students makes only limited use of diagnostic assessments. They would take on a more central role in selecting Tier 3 intensive interventions for students who are performing well below grade level.

Progress-monitoring measures are administered periodically to provide a teacher with feedback as to whether instruction is having the desired effect. The information they provide can be useful in adjusting approaches to instruction in order to improve learning. They answer the perennial educational question "Are they learning what I am teaching?" These measures are frequently the same as those used for screening. This is one example of how the same measure can serve different purposes. Sometimes, though, they are assessments related much more directly to the content of the instruction.

Finally, *outcome measures* help educators judge the effectiveness of instruction on a broader scale. They typically combine the results for many students to measure the achievement of classrooms, schools, districts, and the nation. They include (but are not limited to) the high-stakes tests that so often concern teachers and administrators. We believe that outcome measures serve an important purpose by providing stakeholders with the information they require. Because outcome assessments come near the end of the school year, it is too late to use them to plan instruction, but the results can shed light on how effective the overall instructional program has been and possibly suggest modifications for the next year. Our concern is that teachers make the best use of outcome measures—and that use is very limited. At best, they provide tentative screening information for the upcoming year, although it is frequently too dated to be of much use. Outcome measures do not provide information that is specific enough to guide instructional planning. Moreover, their use for progress monitoring would be cumbersome and inappropriate. What is troubling is that teachers in grades 4 and higher rely mainly on outcome measures (Torgesen & Miller, 2009). It is important to avoid this pitfall by becoming aware of the variety of available assessments and learning how to use them in concert.

Copyright © 2020 The Guilford Press

ORGANIZING FOR ASSESSMENT

The thought of assembling a battery of useful assessments and then coordinating their use may seem daunting: There are so many types of assessments and so many possibilities

for using them. As you will see, however, accomplishing this aim is not difficult. We begin with a few simple guidelines about reading assessment for small-group instruction.

- *Aim for the fewest assessments to answer the questions that are important.* We don't assess for the sake of assessing. An assessment system that is "lean and mean" is far preferable to one that generates a great many data points that no one will use. Because administering assessments takes time, a minimalist system helps to ensure that the time left for instruction will be maximized.

- *Assessments must be coordinated to account for the important aspects of reading.* Identifying a single area of need and directing all available resources toward meeting that need may not be enough to ensure that students become proficient. Too often, students are experiencing multiple problem areas, and it is crucial for teachers to arrive at conclusions concerning particular students' status in word recognition, fluency, and comprehension.

- *All students must be screened.* The fact that students in the upper elementary grades are performing at benchmark levels at the beginning of the year does not mean that no further assessments are needed: "Students must acquire many additional reading skills after third grade in order to be proficient readers in high school" (Torgesen & Miller, 2009, p. 10). This fact requires that we assess even students who are not presently struggling, in order to ensure that they continue to make progress.

- *Formative assessments are the key to successfully using data to guide instruction.* Formative assessments are informal measures that help teachers plan and adjust their instruction. There are three types of formative assessments: (1) those embedded in ongoing classroom instruction, (2) periodic benchmark assessments, and (3) screening and diagnostic assessments (Torgesen & Miller, 2009). All three have a place in our model of differentiated instruction. In curriculum-based measurement, we make a distinction between general outcome assessments, which are good for temporarily classifying students and for gauging their progress from time to time, and skills-based and mastery measures, which are useful for determining whether specific instructional objectives have been attained (Hosp, Hosp, & Howell, 2016). This distinction is very similar to the difference between screening and diagnostic assessments. The former are used to classify and monitor; the latter are short-term in nature and help us plan instruction from cycle to cycle.

Copyright © 2020 The Guilford Press

We turn now to assessments useful in gathering information about the major dimensions of reading. In order to plan appropriate instruction, we need information in three areas. As you will see, however, the assessment burden is light.

Assessing Fluency

Oral reading fluency is the ability to read aloud grade-level text at an appropriate rate and with a high level of accuracy and natural intonation. This definition contains the

three dimensions of fluency that are important to assess: rate, accuracy, and prosody. Most screening in the area of fluency targets the first two of these, and it is common to use a combined metric consisting of words correct per minute (WCPM). Since this metric only counts words read correctly, there is no need to measure the percentage of accuracy separately. Consensus benchmarks for each grade have been established and are presented in Chapter 6. A brief sample of oral reading, typically 1 minute, can provide a quick indicator of whether a particular student is performing below benchmark. Rasinski's (2003) spring benchmark for grade 4 was 118 WCPM and for grade 5 was 128 WCPM. Hasbrouck and Tindal's (2017) norms table reported a spring benchmark of 133 for fourth grade and 146 for fifth grade for students at the 50th percentile. The differences might be attributable to the increased attention paid to the importance of fluency over time. In both sets of norms, there is only a modest increase from grade 4 to grade 5, but the texts students encounter in grade 5 are more challenging.

We consider fluency to be a pivotal proficiency in students' reading development. This is because students who are dysfluent devote too much attention to word recognition and too little to comprehension. Fluency is therefore a prerequisite of comprehension, although it by no means guarantees that comprehension will be adequate. We assess fluency to identify it (or rule it out) as a cause for concern and as a target of instruction. For those students who fall below the fluency benchmark, it is tempting to assume that they are best served by evidence-based instructional approaches for building fluency. However, this is only the case when a full range of word recognition skills has been acquired. Deficits in skill acquisition are one cause of dysfluency, and it is important to determine whether these deficits exist. If they don't, fluency work is indeed appropriate. If they are present, addressing fluency alone is not likely to result in improved proficiency.

Copyright © 2020 The Guilford Press

Assessing Word Recognition

When students reach the upper elementary grades, they should have received instruction on a full array of decoding skills. They will also have encountered many unfamiliar words in text, and they have attempted to apply their skills in decoding those words. By fourth grade, students should possess the skills needed to decode many multisyllabic words. Their ability to do so is grounded in more basic skills. Namely, they should be able to decode nearly every single-syllable word they encounter, and they should be able to recognize many thousands of words on sight without having to decode them consciously. A full diagnostic workup on children who are still struggling with these foundational skills would be time-consuming, to say the least. It would also require considerable expertise. We are not suggesting that fourth- and fifth-grade teachers conduct such detailed assessments. Far from it. When problems in word recognition appear to be causing dysfluency, we recommend only a brief decoding inventory. We have included such an inventory, the Informal Decoding Inventory, in Appendix E.

Figure 4.1 lists the components of our original version of the Informal Decoding Inventory, as it appeared in the first edition of this book. We structured this assessment so that it began with the more basic application of decoding skills in monosyllabic

Part I: Single-Syllable Decoding
• Short Vowels • Consonant Blends and Digraphs • *R*-Controlled Vowel Patterns • Vowel–Consonant–*e* • Vowel Teams
Part II: Multisyllabic Decoding
• Compound Words • Closed Syllables • Open Syllables • VC-*e* Syllables • *R*-Controlled Syllables • Vowel Team Syllables • C-*le* Syllables

FIGURE 4.1. Components of the original Informal Decoding Inventory.

words, followed by their application in multisyllabic words. In the first edition, we recommended that the most efficient way to administer this inventory for fourth- and fifth-grade students was to start with Part II. This practice saved time and avoided unnecessary testing. Only students who struggled with multisyllabic decoding were given Part I.

At this writing, we have simplified our protocol even further. In essence, we have created a screening that replaces the full administration of Part II of the Informal Decoding Inventory. The screening now appears in Appendix E as the Multisyllabic Words subtest, the final subtest in Part I; it is also presented in Figure 4.2 as the Multisyllabic Decoding Screening (which is what we call it when using it separately). The assessment strategy is simple: If a student in grades 4 or 5 is not meeting the beginning-of-year oral reading fluency benchmark, give the 10 items from the multisyllabic subtest. If that student pronounces 8 or more items correctly, you can assume that providing a fluency intervention, without word-level instruction, is appropriate. If a student does not pronounce at least 8 items correctly, administer the vowel teams subtest. That subtest contains both real words and pseudowords, with specific scoring criteria for each. A student who passes the Vowel Teams subtest in Part I, but not the Multisyllabic Decoding subtest, will benefit from a small dose of multisyllabic decoding instruction in addition to fluency work.

Copyright © 2020 The Guilford Press

flannel	submit	cupid	spiky	confide	cascade	varnish	surplus	chowder	approach
							Total		

FIGURE 4.2. The Multisyllabic Decoding Subtest (which appears in the current version of the Informal Decoding Inventory as Multisyllabic Words, the final subtest in Part I; see Appendix E).

For upper elementary students who have not mastered vowel teams, a focus on more basic skills is required. Administering the full version of Part I of the inventory can help verify this need and identify specific instructional targets. However, we do not believe that Tier 2, classroom-based instruction in these basic skills is realistic in fourth and fifth grades. Our reading of current standards requires mastery of all single-syllable decoding skills in the early spring of grade 1. Students in upper elementary grades who need this instruction deserve Tier 3 intervention, not simply small-group instruction in the classroom. This advice should come as good news: It simplifies matters for the fourth- and fifth-grade teacher, and makes an assessment-driven approach practical and easy to manage. Commercial intervention programs in the area of decoding typically include their own assessments. Once informal assessments have indicated the need for such a program, these built-in assessments should be used to guide instruction (Torgesen & Miller, 2009).

Assessing Vocabulary

Although no one disputes the importance of vocabulary knowledge in reading, the problem of assessing that knowledge has proved difficult to solve. The National Reading Panel (2000) identified vocabulary assessment as an especially troublesome area. We have not seen much progress in informal assessments since then. There are dependable normed and standardized screening tests available, to be sure, but they are time-consuming and provide little information that is helpful in planning differentiated reading instruction. Diagnostic tests of vocabulary, in contrast, are nonexistent. This is because a diagnostic test delineates an area into the specific skills a student may lack. In the case of vocabulary, these skills are the equivalent of individual word meanings. That is, every new word is a "skill." Because there is no agreed-upon vocabulary curriculum for each grade, we cannot simply assess a student to determine which words need to be taught. Even if there were such a curriculum, it would contain too many words to make diagnostic assessment feasible.

We do think that strategies for assessing knowledge of taught words are important. Given new standards' linking of reading and writing, we believe that the most important evidence that a student has learned a word is the ability to use that word in a rich sentence-level context. We describe vocabulary instruction in Chapter 7, but we preview it here by sharing the rubric we've developed to assess it. Figure 4.3 attends to both the meaning of the word and the quality of the sentence-level context that a student generated for it.

Assessing Comprehension

Comprehension is unquestionably the most important dimension of reading—the bottom line—and yet assessing comprehension is difficult. However, screening measures, such as the comprehension subtest of a group achievement test, can be useful in determining how a student is performing relative to grade-level expectations. Although we have pursued the design of free curricula for teaching reading and writing, we have not found free

Copyright © 2020 The Guilford Press

	4: Exceeds Standard	3: Meets Standard	2: Progressing	1: Developing
Meaning	Sentence demonstrates clear understanding and is written in a creative way.	Sentence demonstrates clear understanding.	Sentence demonstrates partial understanding.	Sentence does not demonstrate understanding.
Structure	The resulting sentence is compound or complex.	At least 3 questions are answered.	At least 2 questions are answered.	Either 1 question or 0 questions are answered.
Word usage		Word is used correctly.		Word is not used correctly.

FIGURE 4.3. Rubric for rating vocabulary use in context.

screening assessments of comprehension that are valid and reliable. Some schools will have purchased commercial comprehension assessments, but others will not.

Because states typically require outcome assessments yearly beginning at grade 3, spring assessments at the end of that year and the next may be useful as screenings for the fall. Many outcome assessments produce achievement levels in four bands, similar to the NAEP's. For students with spring scores of 3 or 4 (typically equating to meeting or exceeding standards), it is safe to assume adequate comprehension the next fall. For those with scores of 1 or 2, it makes sense to think diagnostically about their comprehension. In order to do that, we use a fluency screening, and, if necessary, the Multisyllabic Decoding subtest.

As is the case for vocabulary, there are no diagnostic tests of comprehension, but the reason is different. Attempts to delineate comprehension into specific skills and strategies have proved fruitless, because assessments of these skills are highly correlated (McKenna & Stahl, 2009). A student who scores high on one skill is likely to score high on others, for example. The best way to diagnose a comprehension problem is to examine the various factors that might contribute to that problem. These include difficulties with word recognition; limited vocabulary and background knowledge; lack of familiarity with various text and sentence structures; and the failure to apply comprehension strategies for specific purposes. We are not suggesting that assessments in each of these areas are needed to implement differentiated reading instruction, but comprehension assessment is nevertheless a part of our approach to differentiated instruction.

Let's begin by considering the two reasons to assess comprehension. One is to determine a student's overall level of proficiency; the other is to gauge the student's understanding of a particular text. These are very different goals, and both are important. Consequently, in our approach, two kinds of comprehension assessments are needed. The first is a screening measure designed to provide an overall level, usually translated into a performance level or Lexile—a metric used to rank students' ability on a scale ranging from beginning reading (a scale score of 200) well into advanced ranges (see *https://lexile.com*). A Lexile is a prediction of 75% comprehension when students are reading on their own. Figure 4.4 demonstrates how a teacher might judge the suitability of a particular book for use with a particular student or with a small group of students. In this case, the

Copyright © 2020 The Guilford Press

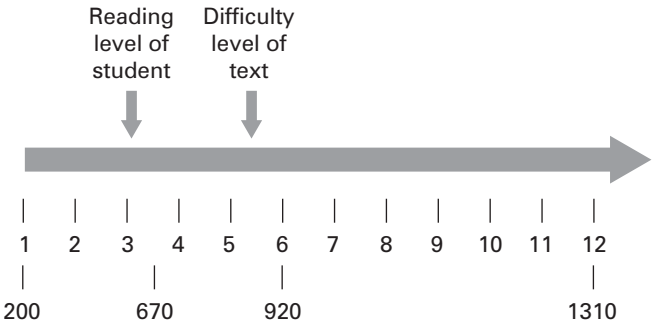

FIGURE 4.4. Using grade levels or Lexiles to judge the match between students and texts.

text is likely to be very challenging without support from a teacher—unless the student is reading about a topic about which he or she has very strong background knowledge.

We believe that comprehension metrics create the illusion of precision. In reality, they are merely estimates that must be weighed, together with teacher judgment. We are also convinced that a precise match between a student and a text is not important; getting reasonably close is good enough. And in some ways, a precise match may not always be desirable. Inspecting this example might prompt one to think that this hypothetical text is too difficult. Keep in mind, however, that in small-group instruction the teacher is in a position to provide considerable support that makes challenging texts appropriate.

The second comprehension assessment useful in small-group differentiated instruction is the day-to-day informal information a teacher derives while interacting with students. This information might come from asking questions or thinking about the questions students ask. If these examples seem imprecise, so be it. Comprehension of a particular text is difficult to reduce to a number or set of numbers even under the best of conditions, and certainly not in the give-and-take of small-group instruction. What is important is for teachers to judge whether comprehension is adequate. If it isn't, adjusting the level of support or switching to an easier text may be required.

As we have designed Tier 1 instruction in *Bookworms K–5 Reading and Writing*, we have designed assessments of comprehension of the grade-level curriculum. Much as we evaluate vocabulary with the vocabulary rubric we have shared in Figure 4.3, we use rubrics to evaluate comprehension. We assign text-based writing prompts to reading every day, and we evaluate one every two weeks. Figure 4.5 shares our rubric for scoring those responses in fourth grade; the fifth-grade rubric has the same format but uses the fifth-grade standards.

Given the imprecision inherent in comprehension measurement, we do not expect a teacher to chart gains in comprehension over the course of a year, and certainly not over a matter of weeks. The tools available for this purpose are simply not very good. Other than informally monitoring students' comprehension of each text they read, nothing more is needed. We focus instead on using assessments to address the underlying factors that impair comprehension and providing a rich diet of texts, high in vocabulary, text structures, and other nutrients.

Copyright © 2020 The Guilford Press

	4: Exceeds Standard	3: Demonstrates Standard	2: Progressing	1: Developing
First, consider accuracy.		Answer is plausible.		Answer is not plausible.
↓	**4: Exceeds Standard**	**3: Demonstrates Standard**	**2: Progressing**	**1: Developing**
Next, consider evidence.	Response refers explicitly to what the text says and includes accurate quotes.	Response refers explicitly to what the text says and includes details and examples from the text.	Response refers explicitly to key details in the text.	Response refers generally to the text or is unrelated to the text.

FIGURE 4.5. Grade 4 rubric for scoring written responses as evidence of comprehension.

Copyright © 2020 The Guilford Press

Assessing the Affective Dimensions of Reading

Affect is a dimension of reading development that is frequently overlooked. It involves how well students like to read, what they like to read, and what they think of themselves as readers. Given the well-documented downward trajectory of reading attitudes and habits, we believe that these factors have a place among the assessments classroom teachers use to improve their understanding of how their students function as readers. We confess that simply documenting that a student harbors a negative attitude or has come to view him- or herself as a poor reader is not of very much help in planning instruction. Other than selecting books that are engaging and accessible, we typically do not make the affective side of reading a primary target. We argue, however, that affect should instead be an indirect target. Supplying an abundance of interesting texts, facilitating students as they engage those texts, and working to build the skills and strategies needed to comprehend them can improve attitudes.

The chief usefulness of assessing affect lies in gauging changes over the course of a school year, not in planning instruction for a group. We suggest that three assessments are sufficient: an interest inventory, an attitude survey, and a self-perception survey. These are group assessments, given at the beginning and end of the year.

An interest inventory is simply a list of topics that might be of interest to students. It could follow a checklist format, allowing students to easily identify those topics about which they might be interested in reading. Some teachers prefer to use graduated responses so that students can indicate their degree of interest. For example, students might be asked to give each topic a "grade." Interest inventories are of two kinds: general and content-specific. A general inventory lists a range of topics and types of fiction. A content inventory lists aspects of a subject area that students might like. A science inventory, for example, might include subtopics of likely appeal (e.g., poisonous snakes, strange phenomena, black holes). Content inventories have utility beyond small-group instruction. The results can be useful in recommending books to students in connection with content area instruction. An example of a general interest inventory is presented in Figure 4.6. Note that it contains a few blanks. The reason is that an interest inventory is essentially

Name _____

Which topics do you like the most? Pretend you're a teacher and give each one of these a grade. Give it an A if you really like it, a B if you like it pretty well, a C if it's just OK, a D if you don't like it, and an F if you can't stand it! If I've missed some topics you really like, please write them on the lines at the bottom of the page.

_____ sports	_____ monsters
_____ animals	_____ horses
_____ magic	_____ detectives
_____ jokes	_____ love
_____ exploring the unknown	_____ famous scientists
_____ sharks	_____ ghosts
_____ camping	_____ other countries
_____ UFOs	_____ dogs
_____ spiders	_____ comic books
_____ the jungle	_____ the ocean
_____ drawing, painting	_____ music
_____ riddles	_____ science fiction
_____ friendship	_____ cats
_____ snakes	_____ families
_____ the wilderness	_____ the desert
_____ fishing	_____ computers
_____ manga	_____ video games

What other topics do you really like? Write them here:

_____

_____

_____

_____

_____

_____

Copyright © 2020 The Guilford Press

FIGURE 4.6. Example of a general interest inventory.

From *Differentiated Literacy Instruction in Grades 4 and 5, Second Edition*, by Sharon Walpole, Michael C. McKenna, Zoi A. Philippakos, and John Z. Strong. Copyright © 2020 The Guilford Press. Permission to photocopy this material is granted to purchasers of this book for personal use or use with students (see copyright page for details). Purchasers can download additional copies of this material (see the box at the end of the table of contents).

a ballot, and every ballot should have a place for write-in candidates. Although you are welcome to duplicate the inventory in Figure 4.6, it is a good idea to create your own, so that you can modify it as needed. For example, you may find that some topics are rarely checked or that topics you overlooked are frequent write-ins. You can edit your inventory accordingly. Finally, you must be able to deliver the goods. It is pointless to include topics for which you have no texts to recommend or to use for small-group work.

A reading attitude survey asks students to respond to statements or questions that are matters of personal judgment and opinion. Questions such as "How do you feel about reading on a rainy Saturday?" are quickly rated on a Likert or pictorial scale. Summing the results provides an overall indicator of whether a student's attitude is positive, negative, or indifferent. A free attitude survey long popular in the upper elementary grades is the Elementary Reading Attitude Survey (ERAS), a pictorial instrument based on the cartoon character Garfield (McKenna & Kear, 1990; McKenna, Kear, & Ellsworth, 1995). It is group-administered and easy to score and interpret. It contains two subscales: one measuring attitude toward academic reading, and another assessing attitude toward recreational reading. The ERAS has excellent psychometric properties and has been used as the basis of numerous research studies. Kear, Coffman, McKenna, and Ambrosio (2000) also developed the Writing Attitude Survey (WAS) to measure students' attitude toward writing. Both may be downloaded free at *www.professorgarfield.org*.

An assessment of self-perception is designed to provide teachers with an idea of how students view themselves as readers. A free instrument specifically designed for grades 4–6 is the Reader Self-Perception Survey (RSPS; Henk & Melnick, 1995). The RSPS assesses four dimensions: (1) progress (how a student views his or her progress in becoming a more proficient reader), (2) observational comparison (how the student compares his or her proficiency with that of peers), (3) social feedback (input the student has received from peers and family about his or her reading), and (4) physiological states (internal feelings that the student experiences during reading, such as comfort or frustration). Like the ERAS and WAS, the RSPS is nationally normed, group-administered, and easy to interpret.

Copyright © 2020 The Guilford Press

A COORDINATED PLAN FOR ASSESSMENT

Now that we have explored the characteristics of various assessment instruments likely to be useful in upper elementary grades, it is time to take stock of what we need and bring the components together in an assessment tool kit. Its contents are listed in Figure 4.7. The left-hand column lists the types of assessments you will need; the right-hand column indicates specific assessments that would be suitable (with room to write in specific instruments that might be available).

Using Assessments Systematically

These tools are enough to accomplish the principal goals of our differentiation model: (1) place students into appropriate small groups, (2) plan instruction targeted to the needs that group members share, and (3) gauge the impact of that instruction on student progress.

Assessments You Need	Assessments You Have or Can Get
Comprehension screening measure from previous year	
Comprehension performance measure	Figure 4.5
Vocabulary performance measure	Figure 4.3
Oral reading fluency screening	
Multisyllabic decoding screening	Multisyllabic Decoding Screening (Figure 4.2; also appears in Appendix E as the Multisyllabic Words subtest in Part I of the Informal Decoding Inventory)
Inventory of decoding skills	Informal Decoding Inventory (Appendix E)
Attitudes toward reading and writing	ERAS and WAS
Self-perception as a reader	RSPS

FIGURE 4.7. Assessment tool kit.

Forming small groups requires systematic use of a few basic assessments from our tool kit. Placement of students into groups is not a precise process. It involves estimation and compromise, but its benefits are considerable. It is useful to think of students with reading difficulties beyond grade 3 as falling into either of two broad categories (Torgesen & Miller, 2009). One includes students who are reasonably fluent but who lack the vocabulary, background knowledge, and comprehension strategies to understand grade-level text. The other includes students who are not fluent and who may lack the more fundamental decoding skills needed to become fluent. Fluency screening can help teachers in grades 4 and 5 quickly decide which category is the better fit for a given student. Students who fall in the second category (usually far fewer in number than the first) require additional informal assessment at the word level. Figure 4.8 represents our grouping process.

The point of giving formative assessments is to follow them with the kind of effective, targeted instruction that the assessments indicate (Torgesen & Miller, 2009). Once groups are populated, such instruction is determined by the focus of the group. As Figure 4.8 illustrates, there are only three types of groups, each with a dual focus: (1) vocabulary and comprehension, (2) fluency and comprehension, and (3) fluency and comprehension with multisyllabic decoding. There is an intended overlap of four areas in these three groups. All three groups have some attention to comprehension. Our top group, vocabulary and comprehension, includes both the students who are known to have strong comprehension and those who may have some comprehension weaknesses but have strong fluency. We can serve both types of students well in the same group. In the next four chapters, we discuss instruction in each of these four areas.

Using Assessments to Regroup

Formative assessments are most useful when they are given periodically, after regular intervals. In our approach to differentiated instruction for earlier grades, we recommend

Copyright © 2020 The Guilford Press

3-week cycles. This length of time not only fits conveniently into marking periods, but it also guarantees that teachers will regularly take stock of student progress and adjust their instruction accordingly. For students in the upper elementary grades, though, a longer group membership is perfectly fine. That longer period allows teachers to use longer books.

After forming the groups, a teacher will embark on instruction for all groups. Each day will bring informal assessment information. Making a few notes about how individual students are responding to instruction will be useful. At the end of the marking period, the teacher is in a position to judge which group will best serve each student's needs going forward. These cycles continue throughout the year, and a combination of formal and informal assessments can determine group membership in a truly flexible way. Remember that students who are experiencing the combined effects of fluency building in Tier 1 and differentiation may improve and be better served by moving up to the vocabulary and comprehension group. Figure 4.9 illustrates this continuing process, beginning with initial screening and proceeding from cycle to cycle through the year, ending in outcome measures (including high-stakes assessments). We note that this year-long model, which is based on the one proposed by Torgesen and Miller (2009), occasionally includes a midyear screening assessment.

Maintaining Records of Assessment

We have deliberately created a Tier 2 instructional system in which assessment plays an important but limited role. Despite its limited nature, however, the information acquired

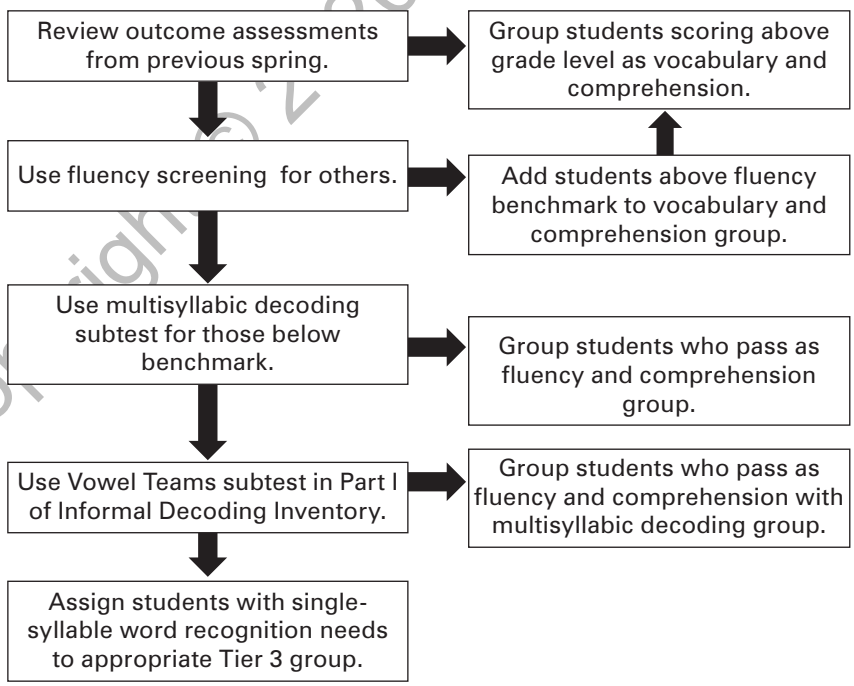

FIGURE 4.8. Using assessments to form groups.

Copyright © 2020 The Guilford Press

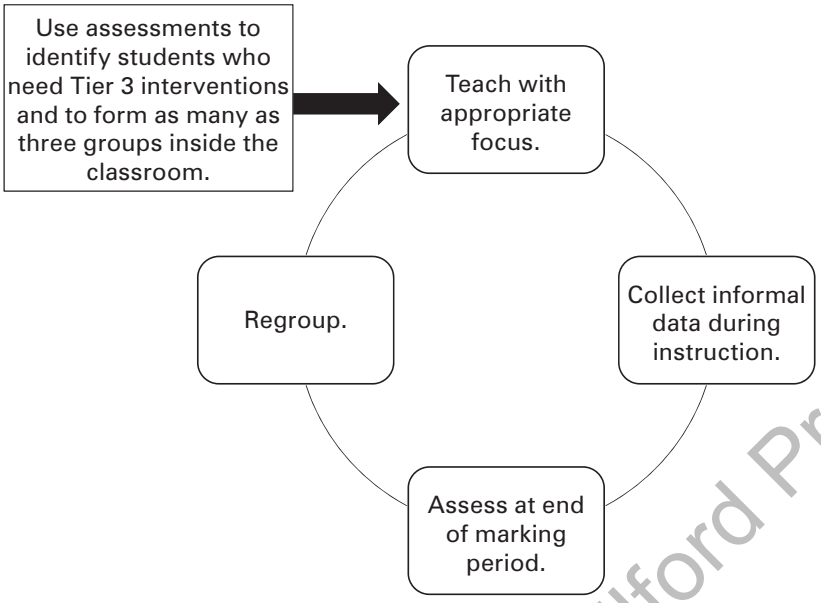

FIGURE 4.9. A year-long plan for using assessments. Based on Torgesen and Miller (2009).

for individual students is surprisingly extensive when we consider it across the course of the school year. This information allows us to track students and to gauge their progress at a glance. One might argue that as long as the general blueprint laid out in Figure 4.9 is followed, there is no pressing need to keep records over time. We disagree and can suggest three persuasive reasons for doing so. First, the decision a teacher must make at the end of each cycle, although based largely on the student's performance during that cycle, can also be affected by the longer history of small-group work. Second, this history will be useful as evidence of the need for Tier 3 intervention for the few students who do not progress. The decision to provide such instruction will probably be reached not by the classroom teacher alone, but in conjunction with a specialist or team, who will benefit from examining the student's history. Finally, maintaining records over time can play a role as part of a larger RTI plan. For example, if Tier 3 intervention proves effective, students can move to a fluency and comprehension group with multisyllabic decoding.

A FINAL WORD

Choosing and using assessments as part of a differentiated instruction plan is not difficult. Those required are few in number and easy to administer. They lead to straightforward group placement, and they provide the information needed to regroup appropriately. They also facilitate a simple long-term record-keeping system that allows a teacher to track the progress of students over time and make decisions about the kinds of instruction that will best meet their needs.

Copyright © 2020 The Guilford Press.
No part of this text may be reproduced, translated, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording, or otherwise, without written permission from the publisher.
Purchase this book now: www.guilford.com/p/walpole4

Guilford Publications
370 Seventh Avenue
New York, NY 10001
212-431-9800
800-365-7006
www.guilford.com