As I prepared to write this article on the impact that Marie Clay has had on assessment, I searched through many books and articles that have recently been written about her contributions to literacy, as well as her own scholarly writing. Interestingly, I found only one article that addressed assessment directly—a chapter in Changing Futures: The Influence of Reading Recovery in the United States (Schmitt, Askew, Fountas, Lyons, & Pinnell, 2005) on observing, assessing, selecting, and monitoring children in Reading Recovery. However, the topic of assessment is embedded within articles that address observation, research methodology, measurement, program evaluation, or the specific tools used in Reading Recovery such as Concepts About Print, Writing Vocabulary, and running records.

I think the lack of articles on assessment in the literature on Reading Recovery may be due, in part, to the fact that Marie Clay has always been a wordsmith—she chose her words carefully. In this case the term assessment has carried with it an association with measuring outcomes, data collected at the end of instruction. I think, in order to establish a clear distinction between her use of assessment in Reading Recovery and the more-typical uses from the field of literacy, she wove assessment carefully into the fabric of her writing. This article addresses how assessment is intentionally embedded within all of our work with children (before we teach, while we teach, after we teach, and as a means to evaluate the effectiveness of Reading Recovery). Marie Clay chose to position our work with children under the umbrella of systematic observation. She chose to name the assessment tools she devised the Observation Survey (2002), and she referred to each tool as a “task.”

Assessment, at its central core, is about decision making. The elegant pebble dropped in the pond—Clay’s focus on observation and decision making—establishes coherence in an intervention for children who are at risk of failing to learn to read.

Rather than a “test,” seemingly simple distinctions, yet these decisions have had astonishing ripple effects.

According to my computer dictionary, assessment means to make judgments about some thing based upon an understanding of the whole situation using a variety of methods to evaluate student performance and attainment. Assessment, at its central core, is about decision making. The elegant pebble dropped in the pond—her focus on observation and decision making—establishes coherence in an intervention for children who are at risk of failing to learn to read. Assessment is infused throughout Reading Recovery; the evidence we gather and judgments we make every day at all levels shape the current and future directions we take. Delineating what we do in the name of assessment (and why) will demonstrate the basic integrity of Reading Recovery and explain Clay’s contributions to innovative assessment practices. She has changed the face of assessment for teachers, school professionals, and literacy educators. The children we have served in our past, we now serve in the present, and we will serve in the future are the beneficiaries of her genius.

In 1999, Linda Gambrell conducted a survey for the National Reading Conference to identify scholars who had most influenced literacy practices across three decades—1970s, 1980s, and 1990s (Gambrell, 2009). While many scholars were named, Marie Clay was the only one identified as having a major influence across all three decades—1970s, 1980s, and 1990s (Gambrell, 2009). While many scholars were named, Marie Clay was the only one identified as having a major influence across all three decades. According to Gambrell, “She changed the way we assess and teach beginning readers. Her influence is seen in the acceptance of observation as an assessment tool. Also, she influenced the shift from the medical model to continuous, naturalistic assessment” (p. 134). I would like to explore the ways in which I see she has influ-
enced assessment, addressing the shift from a medical model to continuous, naturalist assessment, and how Marie Clay has changed the way we assess and teach children. I would also like to stress how assessment is integral to teachers’ learning and to current and future decisions we make to improve Reading Recovery as an effective early intervention.

In 2004, Dr. Clay received the National Reading Conference Distinguished Scholar Lifetime Achievement Award. Her acceptance speech was later reprinted in the fall 2007 special tribute issue of The Journal of Reading Recovery. I was struck by one of her statements as I reread this piece entitled, “Simply By Sailing in a New Direction You Could Enlarge the World.” Clay writes, “I do not study history; I am history” (2007d, p. 8). Born in 1926, she entered college in 1943. The decades Clay lived through were particularly rich.

A Brief View of Assessment Then and Now

Nila Banton Smith (1965, 1986, 2002) wrote a classic text on the history of reading in the United States as a dissertation in 1934. It was recently republished as a special edition. Smith titled the decade from 1900 to 1910 as “Beginning Attention to Reading Disability,” and noted that only 20 research studies were published on reading in this timeframe, with a total of 34 studies in reading reported in the English language arts up until 1910. Edmond Burke Huey’s classic treatise on the reading process was one of those 20 studies (1908). Most of these researchers sought a “specific cause” for “retardation in reading” (Smith, 2002, p. 146). The scientific movement in education was born. Thorndike’s handwriting scale was the first scientific measurement tool published in 1909 and opened up the floodgates to follow. The First World War (1917–18) was a major influence cited as having stimulated this interest in testing, as thousands of soldiers were unable to read printed directions necessary to military life.

Farr (1969) reviewed measurement practices in reading for the International Reading Association. Some of the earliest measurement tools he listed were informal reading inventories such as the Durrell Analysis of Reading Difficulties (developed in 1937) to assess oral and silent reading, listening comprehension, and letter identification, etc. Farr listed the many tests developed in this early time period. The Gates-MacGinitie Reading Test (developed in 1926) with subtests for vocabulary, comprehension, speed and accuracy was one of the most popular. Another early test was Monroe’s Standardized Silent Reading Test (developed in 1919). The testing movement grew rapidly from the 1930s through the 1960s to include a great variety of tools: for screening language disabilities; for assigning remedial treatments; for assessing readiness for reading and diagnosing reading difficulties. Others were for general and reading achievement tests, verbal ability, intelligence tests, picture vocabulary tests, and surveys of study skills and adult basic reading. This movement was heavily influenced by the scientific method being used in the field of medicine, with terms like diagnosis, treatment, correction of disabilities, congenital word blindness, and dysfunction used in the literature. Smith (2002) noted that the testing movement kicked off the period of intensive research investigation from 1910 to 1925, with the peak in 1924 producing a total of 436 research studies. These investigations subsequently led to new methods in the teaching of reading and an increase in professional literature available to reading educators.

A Young Educator Within this Historical Context

Clay mentions her own early experiences administering Stanford-Binet and Wechsler individual intelligence tests (2004, 2007d). She was working as a clinical psychologist during the Second World War. Her M.A. degree program (she studied philosophy, sociology, administration, and psychology) and years of teaching were especially important in her own development. She taught children with low IQs and worked as a psychologist selecting children for special classes. She conducted her first case study in the early 1940s — her first encounter with the specific area of reading difficulties. She noted reading Helen M. Robinson’s Why Pupils Fail in Reading (1946) during her studies, which indicated that difficulties in reading resulted from multiple causative factors. These experiences later helped her design her research and the instruments that, in her words, she “sails by” (Clay, 2007d, p. 8).

A critical event in 1951 changed the course of her life. It also positioned her to do the work we most know her for. She was accepted to go to the University of Minnesota on a Fulbright Scholarship for advanced courses in special education. However, once she was at the university, she found that the professor she was to study with was no longer teaching and there were no courses for her to take! She sought out another program that would accept her, and began to
study developmental psychology at the Institute for Child Study, where she studied with Dr. John E. Anderson.

The faculty at the Institute, like other research laboratories established in North America around 1925, conducted very careful research in developmental psychology (Baltes, Reese, & Nesselroade, 1977, in Clay, 2007d). At the time Dr. Clay was at the Institute in 1951, the field of developmental psychology was well established, having been influenced by the child study movement of the early 1900s (Clay, 1998). As the field of study evolved, these psychologists studied change over time in normal development across the human lifespan. This change from focusing on abnormal learning to studying what occurs in normally developing human psychology expanded Clay's view and further honed her skills.

Three guiding principles from her studies in child development (Baltes, 1983 in Clay 1990, 2007d; Clay 1998) form the basis of the "unusual lens" Marie Clay used that has changed the face of assessment:

1. Describe: In order to understand what occurs in something you study (to describe behaviors, abilities, and processes), you have to descriptively map out what occurs through systematic records of your observations (systematic data collection);

2. Explain: Your observations and analysis (and the theories you derive and test) must reliably explain how and why change takes place (identify the steps of change we should aim for); and,

3. Optimize Development: Enhance development by optimizing opportunities for learning to provide interventions that build bridges for a return to a normal trajectory of progress. This was one of her self-proclaimed "heretical beliefs." (Clay, 2007d, p. 9)

And so, the path that Clay forged that also changed assessment practices was set. She focused on individuals rather than groups, those children who were least likely to succeed, and she explored "what could be done to prevent literacy failures" (Clay, 2007d, p. 9). In order to accomplish these goals, she needed unique instruments that she crafted for her particular purposes. Clay did live the history of assessment, and this vantage point of living through and practicing within both clinical and developmental psychology set the stage for something different in the name of assessment to guide the teaching of children.

Something Different: An Unusual Lens

From the outset, the instruments Clay developed for the purpose of assessment were something different from the norm. The problems she had with traditional measures, and even with measurement research that has fed into current models of "developmental learning," were these:

- Research reports the average achievements of groups, but that average performance does not reflect the actual performance of any one individual;

- Or it reports what the median child of the group was able to do, which also is not a model for all children;

- Or it describes a longitudinal path taken by a particular child, but it is unlikely that any other child will travel precisely that path. (Clay, 1998, p. 89)

In terms of standardized tests, she described them as "indirect ways of observing children's progress...suitable for analyzing the behaviors of groups of children but cannot compare with the observation of learners at work for providing the information needed to design sound instruction" (Clay, 2002, p. 11).

Because Clay chose to study and work with emergent literacy learners and those who were struggling, there were few instruments available to her. Most educational assessments were developed to measure the outcomes of instruction rather than allowing for systematic, reliable observation of learning itself (Clay, 1998). In order to describe, explain, and optimize development for emergent learners, she used a variety of research paradigms for testing and analyzing data — all to make it possible to create sound assessment tools for teachers to use for systematic observations so that assessment could guide their teaching decisions. These instruments had to provide

- a standard task,

- a standard way of administering the task,

- ways of knowing when we can rely on our observations and make valid comparisons, and

- a task that is like a real world task as a guarantee that observations will relate to what the child is likely to do in the real world (for this establishes the validity of the observation). (Clay 2002, p. 12)
Clay argues that the unique combination of a standard task (like a running record of a child’s reading), with guidelines that allow standard administration, and with standard scoring procedures, creates sound measurement conditions—a guarantee of reliability. The two standards any new instrument must meet are reliability and validity. With her strong measurement background, Marie Clay devised unique tasks such as The Record of Oral Language and Biks and Gutches (Clay, Gill, Glynn, McNaughton, & Salmon (1983) and Sand (1972a).

Her groundbreaking research began in 1962 with a simple question: “Can we see the process of learning to read and write going astray early in children’s schooling?” (Ballantyne, 2007, p. 13). In this seminal work, Clay (1966) conducted long-term observations of 100 children across their first year of school. Clay watched as they read and wrote and learned. As she observed, she was faced with the practical problems that researchers face. How can I objectively capture and record what I see as evidence of the phenomenon I am studying? The ingenious ways she devised to capture her observations were the building blocks of the assessment tools she later validated and published. Clay (1989) described the way Concepts About Print (C.A.P.) came to be:

What is CAP? I was watching a 5-year-old try to read a simple caption book. It was 1963 and I was doing the pretests for a research project. I wrote down everything the child did and said. After thinking a great deal about what I saw, I took three caption books of identical format, and devised some tasks that children of this age might do with such a book, and even undid the staples and turned some text and pictures upside down. (p. 268)

This task became part of An Observation Survey of Early Literacy Achievement (Clay, 2002) that was originally published in The Early Detection of Reading Difficulties: A Diagnostic Survey (Clay, 1972b). An Observation Survey is made up of a series of observation tasks to guide teachers in their teaching of young children (Clay, 2002). Each task was devised from Clay’s extensive background testing preschool children and her foundational knowledge of what makes good and bad items within test construction theory (Clay, 1989).

The systematic observation measurement tasks she developed within her research help teachers to observe the child’s use of oral language and how he controls sentence structures and inflections, concepts about print (how print encodes information), the reading of continuous text (taking a running record), letter knowledge, reading vocabulary (words the child knows when reading), writing vocabulary (words the child knows when writing), and hearing and recording sounds in words (both phonemic awareness and the linking of phonemes to letters). (Clay, 2002)

Each observation measurement task was validated to prove its merits as an assessment tool in terms of reliability and validity. Formally, validity evaluates “the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of interpretations and actions based on test scores” (Messick, 1993, p. 13). The evidence that is mounted to prove validity can take many forms. One review (Johnson & Young-Hubbard, 2003) indicates that “Validity evidence provided in An Observation Survey of Early Literacy Achievement includes (a) the extent that the instrument reflects current conceptualizations of literacy, (b) the similarity of students’ scores on the instrument with those on other literacy assessments, (c) the ability of the assessment to predict students’ scores at a later point, and (d) the expected changes in scores as a result of age, instruction, and experiences with literacy.” (p. 43)

The second criterion that must be proven is that of reliability. This basic measurement standard relates to consistency of scores. Will a given student earn the same score if the task is given before more instruction is given? If the scores are consistent, then teachers can rely on the results of the task to make instructional decisions. Another aspect of reliability has to do with comparisons of scores from one task to another that is supposed to be a similar task. Again, Concepts About Print is a great example of this type of reliability. There should be similar scores obtained if a teacher were to use Sand (Clay 1972a), or Stones (Clay, 1979), or Follow Me, Moon (Clay, 2000) as the book the teacher reads to the child as part of the administration of the Concepts about Print task.
As Johnson and Young-Hubbard describe:

Reliability studies examine the consistency of students’ scores (a) on different items within a single assessment (internal consistency), (b) on different test forms that assess the same content (parallel forms), (c) on the same assessment across different occasions (test-retest), and (d) assigned by different observers (inter-rater reliability). The level of consistency is reported as a statistic, referred to as a reliability estimate, that ranges from 0.0 (no consistency) to 1.0 (perfect consistency). (2003, p. 47)

In their evaluation of the evidence provided for the different literacy tasks she devised, Johnson and Young-Hubbard state:

Clay has provided extensive evidence to support the validity of the Observation Survey. Reliability studies indicate that scores from the Observation Survey provide reasonably consistent information to support teachers in making appropriate decisions about students. Overall, the assessment demonstrates many of the qualities required of an effective early literacy instrument. (2003, p. 48)

“For optimal validity and reliability, Clay’s Observation Survey tasks are administered individually, with standardized conditions, directions, and probes to elicit evidence of partial knowledge or strategic operation” (Jones & Smith-Burke, 1999, p. 269). The instruments she developed are revolutionary today as they each offer “an unusual lens” to provide information about individual differences, “an alternative view of progress” (Clay, 2001, p. 46). Clay used the phrase an unusual lens to refer to “any observational or research methodology which gathers detailed data on changes in the literacy behaviours of young children as they learn to read and write continuous texts over a period of time” (Clay, p. 42). Further, she extends this metaphor by saying “It would be an instrument or procedure that could capture how the learner works at learning, and how those ways of working change” (Clay, p. 82).

Dr. Clay ushered in a new era in assessment by using careful research methods. She carefully developed instruments using controlled, direct observation in the natural setting of classrooms (Askew & Watson, 2009). She provided the necessary proofs that are required to validate measurement tools, and she grounded the tools themselves in a solid theoretical knowledge base. Thus, she produced tools that were sensitive to observation and trustworthy.

Assessment and Instructional Decision Making

The assessment tools used within Reading Recovery lessons for children are dynamic assessments that serve a variety of purposes. In my view, these instruments are quite revolutionary for two reasons: (a) because they inform teaching decisions so that teachers can describe, explain, and optimize development every day they teach; and (b) because the assessments were designed to further teachers’ observations of emergent learners on the cutting edge of their learning as part of purposeful activity. As children read books and write messages, they learn from the activity, engage in self-directed learning, and learn by discovery (DeFord, 2007).
There are several key concepts that help to weave assessment, good teaching, and dynamic learning into an effective system to bring about rapid growth: (a) a highly engaged, constructive learner, working within a purposeful activity; (b) a knowledgeable, observant teacher; (c) reliable, sensitive observation tools to monitor change; and (d) a lesson framework and instructional methods designed to optimize student action and teacher decision making. These concepts are central to the success of students in Reading Recovery. This blend of student action, close observation, noticing teachers, skilled decision making, and consistent monitoring to evaluate teaching decisions and steps of change readers need to take are critical ingredients of highly effective teaching crafted to bring about accelerated learning. One of Clay’s theories in action, here, is the power of a self-extending system. The observation tools she devised were intended to focus observations and illuminate what the child is learning so that the teacher can actively support that learning, guiding the child’s tentative first steps to become surer each time they read or write thus optimizing development.

Clay states that within “activity” children construct their own knowledge, and their new learning influences subsequent behavior. Clay called this a “self-improving system, that is, a response system which extends its own capacity,” or “a self-extending system of literacy expertise, as the act of reading expands the range and effectiveness of strategies which the reader can bring to the task, and the size of the practiced response repertoire upon which he can draw” (Clay, 1991, p. 317). To fuel this internal, working system, systematic observation is the assessment tool, and it is utilized in a continuous way within ongoing teaching and learning. In this way, Clay helped to make observation, assessment, and decision making part of an ongoing, fluid, effective teaching/learning system, and influenced the shift from the traditional medical model of assessment to a continuous, naturalistic assessment paradigm.

Assessment That Informs Also Teaches
I have tried to show how Clay’s theories about child development infused her scholarly work, the assessments she devised, and the instructional methods and model she built into Reading Recovery, most recently revised in the book Literacy Lessons Designed for Individuals Part Two: Teaching Procedures (Clay, 2005a). Assessment was critical to her body of scholarly work, as it helped her to see and record important aspects of children’s learning, uncovering such important constructs as the role of reading errors and self-correction in children’s development of skilled reading and writing behaviors (Clay, 1982).

Another important construct her innovative assessment practices uncovered was our need to pay attention to the growing awareness within the child of any feature the child begins to act upon—those first, half right/half wrong responses to any feature of written language that open doors to new learning. “The concept of awareness opens a door to interactions with children that promote literacy learning while avoiding an emphasis on one ‘chosen’ way into literacy” (Clay, 1998, p. 42).

The research she did grew out of the questions she asked, from her basic inquiry stance. She described herself in this way:

I live in a perpetual state of enquiry, finding new questions to ask, then moving on. I do not have ‘a position’ or a safe haven where what is ‘right’ exists. Pragmatism precludes idealism. I search for questions which need answers. What exists in the real world? And how well do our theories explain what exists? Opposing arguments in debates seem to block my search for new solutions, although I have great enthusiasm for brain-clearing discussions. I want to find evidence to convince me of the need for changes in understanding. (Clay, 2001, p.3)

Through her example, she led teachers to inquiry. The professional development program that she devised—as a way to introduce teachers, through observation, to children’s amazing capacity to learn—truly opens new doors and leads them to new horizons. Her model of professional development guides teachers to inquiry. Through conversation, teachers convince themselves of the need for change in their own understanding and teaching practices.

Clay accomplished this by making observation central to teacher professional development (DeFord, 2007). What teachers could see, and learn to understand as significant, could have import for how they observed children and how they acted as teachers. She did this through the use of a one-way glass in which teachers observe lessons in action and talk with their peers about what they see. The teach-
ers are led in this discussion while observing by a more-knowledgeable other who scaffolds their learning, an application of a Vygotskian construct (Vygotsky, 1962).

Vygotsky “applied the concept of a zone of proximal development to both instruction and diagnosis” (Clay & Cazden, 1990, p. 220). One of his well-known quotes applies to children as well as teachers: “The only good kind of instruction is that which marches ahead of development and leads it; it must be aimed not so much at the ripe as at the ripening function” (Vygotsky, 1962, p. 104). The genius of Marie Clay was to juxtapose teachers in an interactive, conversational setting in which their observations, along with the guidance they receive from peers and their instructional leader, could mold their emerging new views of children as learners. In this setting, several simultaneous interactions occur:

The child’s reading and writing gives the teacher feedback about teaching decisions and informs the nature of the next decisions and actions taken. On the other side of the glass, teacher colleagues are engaged in constructive activity—the observation and analysis of the lesson. They talk about the processing they see, ask questions, and consider different explanations in light of the teacher-child interaction. (DeFord, 2007, p. 101)

Jones and Smith-Burke (1999) note, “The continuous interplay between what teachers and children are doing and thinking facilitates teachers’ construction of personal theories about each child that are grounded in observational data” (p. 273). This dialogic process engages teachers in a learning process that is similar to the self-extending system Clay ascribes to children’s learning. Teacher’s learning is shaped through continuous application of observation, conversation about what they see and think (which is a form of constructive activity), feedback on teaching decisions, and ongoing assessment.

The value of this behind-the-glass conversation is discussed in greater detail by Lyons, Pinnell, and DeFord (1993) who suggested six principles of learning that are foundational to this professional development model: (a) a constructive view of learning, (b) using language to learn, (c) valuing tentativeness, (d) flexibility, (e) creating supportive networks to extend learning, and (f) ongoing learning. Key to this enterprise is an understanding that as teachers acquire knowledge gained through observation and ongoing assessment, they will begin to use theories to guide their instructional decisions. This shift to a theory-based decision model of teaching is what brings about change for teachers and children. As one teacher wrote, “In the intensity of teaching, we can forget how much focused observation—not only of the student, but of ourselves—can further develop our theories about our students’ learning. In reflecting about how we as individuals observe, we can know ourselves and our students better” (O’Leary, 2009, p. 37).

**Assessment for Measuring Outcomes**

I’ve described the reliable and valid, finely tuned instrumentation that captures observations and change in early literacy behaviors, and how assessment is used in each interactive lesson in a child’s instructional program. I’ve discussed how assessment serves teachers’ decision making and learning. The other significant contribution within Clay’s vision and model is at the system and organizational level: the implementation of this program and how the measurement of outcomes informs decision making at the program level.

Anthony Bryk offered technical advice for research on Reading Recovery at The Ohio State University, a project that was funded by the McArthur Foundation. He co-authored the research report that proved the effectiveness of Reading Recovery (Pinnell, Lyons, DeFord, Bryk, & Seltzer, 1994). He is now president of The Carnegie Foundation for the Advancement of Teaching. In a recent article he stated,

Looking back now, I see a program whose organizational design was way ahead of its time. Much of what many now view as core principles for advancing more-ambitious instruction at scale existed two decades ago in the Reading Recovery initiative as it built on and extended the extraordinary developments ‘down under’ by Marie Clay and her colleagues. (Bryk, 2009, p. 17)

Bryk posited three big ideas that he believes work together in this avant-garde organizational design: (a) a detailed, supported instructional system; (b) a clinically based professional education program; and (c) an organized professional learning association seeking continuous improvement by routinely gathering and analyzing data on student learning.

Part of what he saw in this retrospective analysis of Clay’s scholarly contributions was the instrumentation for assessing student learning. “The
In our efforts to maintain effectiveness and improve our ability to serve the lowest-achieving students in first grade, the data we routinely collect are the backbone of evidence we use — not only data collected on individual children served, but in questions we ask to evaluate program effectiveness.

regular collection of running records and the Observation Survey data provides a common evidence base to inform both individual teachers’ day-to-day instruction as well as collective efforts at continuous improvement of the instructional system over time” (Bryk, 2009, p. 18).

In our efforts to maintain effectiveness and improve our ability to serve the lowest-achieving students in first grade, the data we routinely collect are the backbone of evidence we use — not only data collected on individual children served, but in questions we ask to evaluate program effectiveness: ways to serve more children, ways to improve teaching, ways to improve school performance, ways to consider innovations that may offer new initiatives that could extend our own knowledge, etc. This drive to improve at the system level is another iteration of a self-extending or self-improving system. In a speech Dr. Clay delivered in February 2005 (published that year in The Journal of Reading Recovery) she reviewed a 2000 speech by P. David Pearson, in which he outlined factors that contributed to Reading Recovery’s success. Her response to him was this:

These are important and helpful observations from a respected reading authority. Perhaps due to such positive reviews, people expect me to be very satisfied, to be walking around beaming! They are puzzled when I say I am not satisfied. My dissatisfaction stems from knowing that for every one million children we are able to serve, there are six or seven more in the same age group who need our help but who have no access to it. I also know that the particular time in a child’s life that we work with is the only time teachers can get complete recovery. Any time later will be less effective. So, I remain unsatisfied and concerned. (Clay, 2005b, p. 2)

In this talk Clay addressed four issues that require our vigilance if we are to learn, grow, and change as Bryk termed “a professional learning association.” These issues are central to our mission: 1) the need to expand to serve more children; 2) the need to rethink our explanations, 3) the need to explain what’s special about instruction that is individually designed and individually delivered, and 4) the need to address implementation concerns” (Clay, 2005b, p. 2).

In order to reach the goals she set—to expand to serve more children, especially if schools already serve and discontinue the lessons of the lowest 20% of at-risk first graders; or to expand Reading Recovery at the district level and not depend on federal monies to provide needed intervention services; or to create deeper understanding and better explanations about key aspects in the teaching of individually designed lessons. Clay suggested that the top priorities must be the integrity of the teaching and implementation decisions we make and careful monitoring of student outcomes.

This inquiry-based, self-improving stance was applauded in Bryk’s 2009 article. He iterated that a core principle that has guided the development of Reading Recovery is that “you cannot improve practice at scale unless you measure both its core processes and outcomes” (p. 19). We have maintained empirical evidence across 25 years of research, including international studies (see Askew and Watson, 2009, for more information). As Bryk noted, our working theories, instructional processes, organizing routines, materials, and cultural norms are foundational to this empirical system. These interact to affect student outcomes. “A key strength of Reading Recovery in its inception and in its organizational life today is in its openness to continuous learning and in being prepared to challenge its working theory based on new data” (p. 19). He goes on to say:

Ongoing empirical evidence about efficacy in action is the only sure assurance that we ‘know what works.’ Maintaining student learning in the first position, and constantly scrutinizing what is and is not happening for different children who are being educated under different sets of circumstances and instructional contexts, is the essential check and balance for a professional community.
Concluding Thoughts

I hope I have highlighted some of the amazing contributions Clay’s unusual lens on assessment has made to teachers and literacy professionals worldwide, and the service she has extended to the lowest-achieving, young children we teach. I concur with Peter Johnston who said:

Marie has pushed us to maintain coherence among theories of development, literacy, learning, teaching, and teacher education—a coherence based on meaning-directed, self-extending systems. In the process, Marie’s work has shown her enormous respect for children and for teachers, which in itself is a revolutionary contribution. (2005, p. 4)

Marie Clay lived within a world context that sought to dissect, separate out, and control intervening variables through measurement practices using randomized trials, controlled assignment, stepwise regressions, averaged data of group performances, and a medical model that targeted abnormal behavior or explanations of ‘dysfunctions’ that were based upon single-factor diagnoses for reading disabilities. In contrast to the prevailing attitudes, Marie Clay persisted in applying naturalistic, descriptive techniques that mapped the progress of individual children to describe complex interactions and learning puzzles. She then followed up by employing a variety of research methodologies that maintained standards of rigorous scientific methods that are highly respected by the larger research community (see Askew, 2009; Jones & Smith-Burke, 1999 for further information).

Clay’s use of metaphors to describe her theories and scholarly endeavors always amazed me. But one will always stay with me, from a poem about the discovery of New Zealand in 1642. When I recite this line, it gives me new direction and a far-reaching vision to navigate by, no matter what I do or where I am: “Simply by sailing in a new direction you could enlarge the world” (Clay, 2004, p. 60). She did that with such impeccable style, and in my view, changed the face of assessment for all of us.

References


the National Reading Conference (pp. 60–66). Oak Creek, WI: National Reading Conference.


