The assessment of children's reading progress has long been of interest to teachers, researchers, and parents. The purposes for reading assessment include comparing one child's progress to that of his peers, screening students for special assistance, measuring an individual's progress over a period of time, diagnosing particular areas of strength or weakness, using information for decisions about instruction, and determining placement within a reading program or special facility. There have been many different approaches to reading assessment based partly upon these differing purposes, but also upon the conception of reading development held by the test designer or user.

Reading Miscue analysis is a major whole language test designed to assess the strategies that children use in their reading. Goodman and his colleagues in the 1960's were interested in the processes occurring during reading, and believed that miscues (any departure from the text by the reader) could provide a picture of the underlying cognitive processes. He used the term miscue, rather than error, reflecting the view that a departure from the text is not necessarily erroneous (Goodman, 1979). Readers' miscues include substitutions of the written word with another, additions, omissions, and alterations to the word sequence.

Initially, he developed a Taxonomy (Goodman, 1969) which detailed 28 different types of miscues. Established initially for research purposes, its unwieldiness and a desire to broaden its usage led Yetta Goodman and Carolyn Burke in 1972 to develop a briefer version comprising nine questions to be asked about each miscue - a simpler system that they believed would become a useful and manageable tool for clinics and for teachers in the school system. The authors were less interested in traditional quantitative measures such as reading accuracy or reading rate, and considered that their qualitative approach provided more fine-grained and relevant information than did other approaches to reading assessment. In the Reading Miscue Inventory (RMI) a student's incorrect response, when compared to the written word, may display a dialect variation, an intonation shift, graphic similarity, sound similarity, grammatical similarity, syntactic acceptability, semantic acceptability, meaning change, and self-correction with semantic acceptability to the text word. An inventory of a child's miscues involves selecting text of sufficient length and difficulty to ensure that a child will make at least 25 errors, tape-recording the oral reading, and assigning the child's miscues to one or more of the nine categories. A further step requires a retelling of the story by the student as a comprehension check. The RMI requires about 20-40 minutes with each individual child, and a further hour for analysis. Through miscue analysis, the authors argued, teachers can better monitor a child's progress along the path to reading success, and identify the strengths and needs of students. Depending on the prevalence and type of miscue, teachers may decide whether any intervention is required and also its focus.

The value of any assessment tool depends upon the importance of the quality to be measured and the capacity of the tool to perform its task. For example, the measurement of hair length does not provide important information about reading. Despite the fact that hair length can be measured quite accurately using appropriate instruments, it has neither theoretical relevance to reading nor does it correlate even moderately with reading development when both hair length and reading are assessed across the population. Thus, a consideration of the RMI involves two questions, each of which must be answered in the affirmative for the Inventory to be useful: Are the qualities that the instrument purports to measure significant as indicators of reading progress given the current knowledge about reading and its development? Is the instrument a valid and reliable indicator of the presence or absence of the targeted qualities?
Problems with the theoretical basis for the RMI The first question for the RMI involves its theoretical relevance to reading development and of skilled reading? This question is crucial because miscue analysis is predicated upon the whole language conception of reading, and hence stands or falls on the validity of this conception. The significance of any reading errors is thus superimposed on the reading behaviour through the adoption of the whole language view of reading and skilled reading. This question is crucial because miscue analysis is predicated upon the whole language conception of reading, and hence stands or falls on the validity of this conception. The whole language philosophy conceptualises reading development as the gradual integration of three cueing mechanisms (semantic, syntactic, and graphophonic), although the graphophonic system is considered a lesser contributor - even potentially disruptive if over-relied upon by readers (Weaver, 1988). Reading should entail as little emphasis on each word's letter construction as possible. Rather, it is perceived as a process of ongoing prediction of target-words based primarily upon semantic and syntactic cues, followed by confirmation that the chosen word is consistent with the context (and possibly the target word's initial letters). "In turn (the reader's) sense of syntactic structure and meaning makes it possible to predict the graphic input so he is largely selective, sampling the print to confirm his prediction" (Goodman, 1973, p.9).

Consistent with this view of skilled reading, the Reading Miscue Inventory is concerned largely with errors that cause a loss of meaning, the number of errors being less important than their immediate impact on comprehension (Weaver, 1988). There are differences in the acceptability of various miscues. Good miscues maintain meaning and are viewed as an indication that the student is using meaning to drive the reading process, and hence, is on the correct path. Bad miscues are those that alter meaning. Whether the word the student reads corresponds to the written word may not be important in this conception. "Accuracy, correctly naming or identifying each word or word part in a graphic sequence, is not necessary for effective reading since the reader can get the meaning without accurate word identification" (Goodman, 1974, p. 826).

More recent research has demonstrated that this assertion is incorrect. Good readers, though more sensitive to context cues to elicit the meaning of unfamiliar words, do not need to use context to decode unknown words (Tunmer & Hoover, 1993). At best, even good readers can guess words only one time in every four, and then only with fairly predictable words (Gough, 1993). They soon learn that word structure more reliably supplies the word's pronunciation than does context; unfortunately, it is poor readers who are more likely to invest attention on such context guesswork (Nicholson, 1991). The error made by whole language theorists is to confuse the desired outcome of reading instruction - a capacity to grasp the meaning of a text - with the means of achieving that end. In order to comprehend meaning, the student must first learn to understand the code (Foorman, 1995).

A teacher using the RMI will examine the nature of the errors the student has made in the chosen passage. Consider this text and a reader's response, substituting pony for horse:

Child # 1: pony The man rode his horse to town.

Asking the nine questions reveals that the miscue (compared with the target word) has grammatical similarity, syntactic acceptability, semantic acceptability, does not change meaning, and the miscue does not involve dialect variation, an intonation shift, graphic similarity, sound similarity, or self-correction. Such an error is considered an acceptable miscue. Reading pony for horse is indicative of the student using contextual cues appropriately and a signal for satisfaction about reading progress. The teacher would be content with this error, as meaning has been more or less preserved. "Often substitutions of words like a for the, by for at, in for into, do not cause a change in meaning. ... substitutions like daddy for father, James for Jimmy ... are generally produced by proficient readers and are not reading problems" (Goodman & Burke, 1972, pp. 101-102).

According to the whole language conception of skilled reading, students must make many miscues during the progressive integration of the cueing systems in order for reading to develop. It is expected that these errors are not necessarily a cause for intervention but simply a sign of a reader prepared to take risks. Any corrective feedback regarding errors is risky as it may jeopardise the student's willingness for risk-taking."
... if these resulting miscues preserve the essential meaning of the text, or if they fail to fit with the following context but are subsequently corrected by the reader, then the teacher has little or no reason for concern” (Weaver, 1988, p.325).

Suppose another student reads house for horse:

Child #2: house The man rode his horse to town.

Asking the same nine questions reveals that the miscue (compared with the target word) has graphic similarity, some degree of sound similarity, grammatical similarity, syntactic acceptability, and the miscue does not involve dialect variation, an intonation shift. Further, it does not include self-correction, is not a semantically acceptable change, and the miscue creates meaning change. This response is considered an unacceptable miscue because it changes the meaning. "Proficient readers resort to an intensive graphophonic analysis of a word only when the use of the syntactic and semantic systems does not yield enough information to support selective use of the graphophonic system” (Goodman, et al., 1987, p.26). Despite the closer graphemic similarity of the response house to the target word, children who make errors based on graphemic similarity, such as house for horse, are considered problematic and over-reliant on phonic cues. Whole language theorists argue that good readers' miscues display less grapho-phonemic similarity to the target word than do those of poor readers (Weaver, 1988), and readers-in-training should do likewise.

Thus, the remedy the teacher would choose for child #2 would be to encourage him to rely more on context and to look less at letter patterns. However, according to recent research, this remedy is more likely to result in poorer reading than in better reading. Adams (1991) argues that to improve this child's reading, he should be encouraged to look more closely at the letters, the opposite of that recommended in the RMI. Adams found that good readers' miscues display more grapho-phonemic similarity to the target word than do those of struggling readers. In fact, most nascent readers' miscues shift over time, from early errors based upon contextual similarity to those based upon graphemic similarity; however, this shift is now recognised as functional and a characteristic of progress. The student's dawning understanding of the pre-eminence of a word's graphemic structure encourages close visual inspection of words, a strategy that accelerates the progressive internalisation of unfamiliar spelling patterns, that is, it leads ultimately to whole-word recognition. That some teachers may unwittingly subvert this process, with well-meaning but unhelpful advice to readers, is an unfortunate outcome.

According to current knowledge, the house response is a preferable error to the pony substitution. It may be a sign that the student is in the process of acquiring the alphabetic principle; however, corrective feedback should be provided as house is an erroneous response. Through the error correction, the student's attention is directed toward the letters in the written word and the sound usually made by the or combination. The response recommended to teachers through the RMI - that of directing the students' attention away from the letters in the word to that which can be predicted and which makes sense - provides an alarmingly unstable and counter-productive rule for students.

Child #1 is likely to be in greater need of instruction that directs his attention to the letters in the words. Child #1 might equally have substituted bicycle for horse. The substitution makes sense but is far from that which the author intended. The child whose primary decoding strategy is driven by semantic and syntactic similarity is unaware that bicycle bears no graphemic similarity to horse. The instructional message to the student is that, despite the errors directly attributable to the strategy of guessing, the strategy is the appropriate one. The student is encouraged to continue using a strategy that is unhelpful, and is dissuaded from attending to the major cue that would improve his reading - the word's structure. Regardless of the type of miscue, students who make errors need to focus on the letters in the word to improve their reading.

Self-corrections are errors that are corrected without another's intervention, usually because the word uttered does not fit in the context of the sentence. Within the whole language framework, self-corrections are a clear and pleasing sign that meaning and syntactic cues are being integrated into the reader's strategies. Clay (1969, cited in Share, 1990) asserted that good readers self-corrected errors at a higher rate
than did poor readers. She considered high rates were indicative of good text-cue integration, which in turn was a measure of reading progress. The significance of self-correction has been questioned by Share (1990), and Thompson (1981, cited in Share, 1990). They found that self-correction rates are confounded with text difficulty. When text difficulty was controlled in reading level-matched designs, the rates of self-correction became similar among good and poor readers. That is, when text is very difficult everyone is more likely to make errors and increase their rate of self-correction. Hence, an increased rate of self-correction can be interpreted as simply indicative of excessively difficult text rather than as reflecting reading progress. This interpretation based on difficulty levels also raises concerns of unreliability in the assessment of self-correction rates. The conclusion that there is no direct support for self-correction as a marker or determinant of reading progress makes the activity of recording such ratings for students of questionable value.

How does the view of reading underpinning the RMI sit with research findings regarding the reading process and its development?

This view of skilled reading, which comes from Goodman (1967) and Smith (1978), has been rejected by the scientific community (Adams, 1990; Ehri, 1986; Goswami & Bryant, 1990; Gough, Ehri & Treiman, 1992; Just & Carpenter, 1987; Perfetti, 1985; Rayner & Pollatsek, 1989; Rieben & Perfetti, 1991; Stanovich, 1986, 1991; Vellutino, 1991). Skilled reading is not sampling features of the text on the run, it is not a psycholinguistic guessing game, and it is not incidentally visual. Rather, research has shown that 'skilled readers process virtually all the words they encounter in connected text, and typically, all of the letters in those words' (Vellutino, 1991, p. 82). Research further indicates that skilled readers are sufficiently fast and accurate at recognising words in text to make reliance on contextual information unnecessary (Perfetti, 1985). (Tunmer & Hoover, 1993, p. 167)

The findings of individual researchers and such syntheses as provided above have been formalised through the National Institute of Child Health and Human Development (NICHD). In 1985, the Health Research Extension Act directed the NICHD to coordinate research on reading disability and learning disability such that results of research would meet a number of criteria regarding scientific rigour. The intention was to define research characteristics that would ultimately lead to methodologically unassailable findings and benchmarks of consensual knowledge. More than 100 researchers in numerous sites across the USA are involved in this cooperative multi-disciplinary research employing large scale longitudinal studies, careful sampling, and replication of findings with the view of integrating their research efforts. A summary of the findings is provided by the director, G.R. Lyon (1996): The ability to read fluently for meaning depends primarily on rapid, automatic decoding and recognition at the level of the single word. The basis of reading deficits (phonological processing) should provide the focus for intervention. Efforts should be directed at explicitly and systematically teaching the connection between phonological rules and the written word. A phonics emphasis provides advantages for disabled readers over a whole language approach.

The NICHD research summary has been very influential, even at a political level. Recently, the US federal Reading Excellence Act was passed by both houses, is currently in committee, and is expected to be enacted during this year. This legislation ensures that all reading programs eligible for federal support in future will be based on reliable and replicable research. Part of the definition of reading included in the Bill provides a clear indication of consensus concerning the incompleteness of the whole language view of reading. Reading is described as the process of comprehending the meaning of written text by [depending] on the ability to use phonics skills (i.e., knowledge of letters and sounds) to decode printed words quickly and effortlessly both silently and aloud.

In a similar vein, the British National Literacy Strategy (1998) has recently been released to all primary schools, requiring them to abandon the current Whole Language approach to reading. Components of the former system, such as reliance on context clues to aid word reading, are discredited in the Strategy, and schools are directed to introduce explicit phonics instruction from the earliest stages of reading.
The RMI was designed to provide a "window on the reading process" (Goodman, 1973, p. 5); however, the analogy with a window is a misleading one as it implies a direct and transparent medium. The picture of reading obtained through the RMI involves an interpretation of what is viewed through this window. What is really displayed by a student is reading behaviour (words, sentences) - the subsequent analysis of miscues involves making inferences about unobservable processes based upon assumptions about the reading process. With this instrument, the picture is coloured by a discredited conception of reading.

An important rationale for the choice of an assessment device resides its capacity to inform intervention (Goyen, 1992). The RMI Manual, however, provides few strategies for corrective intervention - perhaps because miscue analysis was not originally developed to inform intervention. Alternatively, it may relate to the whole language view that reading progress is natural in a strongly literate environment. "Learning is continuous, spontaneous, and effortless, requiring no particular attention, conscious motivation, or specific reinforcement" (Smith, 1992, p.432). The typical global recommendation from the RMI for students with reading difficulties involves prompting the increased use of psycholinguistic guessing - "The reading strategies of sampling, predicting, and confirming are the same for all readers. ... non-proficient readers ... need to be invited to do what proficient readers do, their attention drawn to inferential strategies" (Goodman, Watson, & Burke, 1987, p. 170). Additionally, students are expected to learn about reading through their mistakes rather than through instruction from teachers; hence, the reticence of the RMI developers toward explicit intervention strategies is understandable, even if unhelpful to students.

One implication of the current understanding of the reading process is that the qualitative analysis of reading errors is largely irrelevant to instructional planning. Decoding errors of whatever type are best addressed at the level of decoding instruction (Lyon, 1996). Thus, the student who makes errors due to reliance on contextual strategies and the student who makes errors based on inadequate graphophonic analysis each requires decoding instruction and practice, sufficient to enable effortless reading at the appropriate level of text difficulty. Psychometric studies have demonstrated that it is decoding ability that predicts children's capacity for word identification and comprehension. Measures of semantic and syntactic ability as assessed in the RMI are not strongly correlated with word identification or passage comprehension (Vellutino, 1991).

Issues of validity and reliability Modern research has indicated that the RMI does not provide important information regarding reading, and hence is of largely historical interest. However, even if its foci were of interest, there are other difficulties that create problems for its use. An assumption implicit in miscue inventories is that oral reading reflects similar processes to those involved in silent reading, and hence errors detected while students are reading aloud are representative of errors in their silent reading. However, even Goodman (1976) expressed suspicions about the usefulness of the results of oral reading assessment " ... poor oral reading performance may reflect a high degree of reading competence rather than a lack of such competence" (p.489). If poor oral reading can be interpreted so diversely, and may be simply an artifact of the assessment, its value is dramatically compromised (at least insofar as its implications for silent reading ability).

An additional problem for the Reading Miscue Inventory is its inadequacy as a psychometric instrument (Allington, 1984). Leu's (1982) review of oral reading error analyses highlighted serious problems of unreliability. Unreliability in an assessment means that the same tool can provide differing results on different occasions, or with different texts, or with different examiners, without any change in the student's capacity. The unreliability problems arise from: (a) vague definitions of the boundaries of the error categories. Determining when meaning has been essentially preserved may produce different decisions from different teachers for the same miscue, (b) an absence of theoretical justification for the categories, (c) a failure to allow for the effects of passage difficulty. When passage difficulty is controlled (i.e., similar error rates), reliance on context occurs at least as much for less skilled as for skilled readers (Allington & Fleming, 1978, Batey & Sonnenschein, 1981; Biemiller, 1970, 1979; Cohen, 1974-5; Coomber, 1972; Harding, 1984; Juel, 1980; Lesgold & Resnick, 1982; Perfetti & Roth, 1981; Richardson, Di Benedetto, & Adler, 1982; Weber, 1970; Whaley & Kibby, 1981; cited in Stanovich, 1986), (d) the ambiguity resulting when categorising multiple-source errors.
Hood (1982) noted that other text characteristics (besides difficulty) also influenced the type of error produced by readers. Wiederholt and Bryant (1987) further point out that miscues are influenced by the reading instruction the students have received, the student age, the writing style, student familiarity with the text, and the stated purpose of the reading task (e.g., reading for expression compared to reading so as to answer comprehension questions). Such contamination of results inevitably leads to inconsistent diagnosis and similarly inconsistent instructional implications. The findings of any individual assessment cannot be relied upon to provide information about the habitual strategies used by a reader, and thus fail the basic reliability requirement of an assessment instrument. Given the difficulty in separating these various potential causes of miscues, it is difficult to accept Goodman's characterisation of the miscue analysis as providing a clear insight into the student's cognitive processes.

The Reading Miscue Inventory has had considerable influence in instructional texts and in classrooms (Allington, 1984), and remains influential among Whole Language theorists and teachers (Weaver, 1988). A revised version - RMI: Alternative Procedures (Goodman, Watson, & Burke, 1987) offers four analysis options of varying complexity for classroom use. The rationale is unchanged "... it is best to avoid the common sense notion that what the reader was supposed to have read was printed in the text" (Goodman et al., 1987, p.60) - and the Alternative Procedures are subject to the same criticisms as earlier versions. Although the RMI has been a very popular test, many teachers (for example, in Reading Recovery) have been trained to use an informal procedure of maintaining "running records" (Clay, 1985) with their students, a procedure that provides similar information on types of errors and self-correction rates, and that is based on a similarly flawed conception of reading.

It should be noted that this critique does not necessarily imply that the qualitative analysis of readers' errors is valueless. However, it is essential for any proposed analysis that the rationale for error categories should be well grounded in knowledge about the reading process, the function of the analysis should be clearly explicated, and the instrument have acceptable psychometric properties. It should also be recognised that such qualitative analyses may be primarily of research interest - but not of immediate use in informing intervention. The authors of the RMI make the claim that their assessment is authentic because it makes use of "real" literature for the assessment task. However, even authentic assessments should meet the requirement of relevance and trustworthiness. Given the problems with the theory, design, and implications of the Reading Miscue Inventory (and its derivatives), its widespread continued acceptance in the education community is difficult to fathom.

A more valuable approach to assessment is one grounded in up-to-date knowledge of skilled reading and its development. Given that skilled readers process almost every letter of every word, and they read fluently and accurately, then we should assess how well we teach students to develop the range of salient skills. We want students to appreciate: How words are composed of phonemes, How words can be deconstructed into phonemes, How phonemes can be blended to manufacture words, How printed letters have sound values (individually and in clusters), How decoding operates, How adequate practice at decoding leads to whole word reading, How fluent reading releases attention from the mechanics of reading to the processes of comprehension, How the use of comprehension strategies can increase our capacity to become more sophisticated readers.

We require tools that provide the appropriate measure for the reader's stage of development. For example, phonemic awareness is strongly predictive of reading success for beginning readers (Adams, 1990), and the early use of such measures can both aid in determining the intensity of teaching required by individuals to ensure their success, and can also reduce the disheartening effects of failure at this critical task. Additionally, instruction in phonemic awareness has been demonstrated to be beneficial to students at risk of reading failure (Adams, 1990; Stanovich, 1986). Thus, the phonemic awareness assessment provides information about students in an important reading pre-skill, and it also guides instruction. Other measures of letter-sound knowledge, blending, and segmenting can assess progress in the decoding development phase, providing information about the adequacy of teaching the student has received and the level of practice required for each individual to master these skills. Assessment of sight words can provide information about the development of orthographic images of words, the stage beyond decoding indicative of progress towards efficient, skilled reading (Ehri, 1995). A schedule of reading-rate assessment is
important (Samuels, Schermer, & Reinking, 1992; Slocum, Street, & Gilberts, 1995) as it is indicative of the reader's growing fluency. A secondary benefit is that rate assessment is simultaneously an intervention as it provides an impetus for students to increase their rate, thereby indirectly aiding comprehension. The regular, frequent measurement of reading rate and accuracy using literature of known and appropriate difficulty levels becomes critical for students who have mastered the earlier stages described above. It guides the teacher in the continuing decisions regarding instruction and student practice needed to ensure the student reaches that stage of automatic, accurate reading described by Ehri (1995) as the "consolidated alphabetic phase" (p. 121). Share (1995) sees the alphabetic period as crucial, and he developed a self-teaching hypothesis in which "... each successful decoding encounter with an unfamiliar word provides an opportunity to acquire the word specific orthographic information that is the foundation of skilled word recognition and spelling" (Share & Stanovich, 1995, p. 18). This gradual "lexicalization" (p. 18) occurs through repeated opportunities to use letter-sound correspondences for decoding. The strategy is used with less frequency as the range of familiar word patterns increases, through a "self-teaching" (Share, 1995, p. 155) mechanism. The phonological recoding strategy remains useful for decoding unfamiliar words - and of course, our language has many low frequency words.

A different conception of reading underpins this approach to assessment compared to that of Reading Miscue analysis; a further major difference concerns the responsibility assigned to the teacher for a student's reading progress. The whole language teacher believes that reading is as natural as learning to speak (Smith, 1992), and acts largely as a facilitator for students, observing the inevitable success engendered by immersion in authentic literature. In contrast, the focus described above includes an emphasis on the importance of quality instruction in ensuring success (particularly for at-risk students). It has been noted that written language is a human invention, and reading should not be viewed as a natural extension of a biological imperative - learning to speak (Liberman & Liberman, 1990). Although it is the learner whose progress is being assessed, the prime purpose is to assess the instructional adequacy of the learner's educational environment. It is a focus on adapting instruction to the progress of the learner that provides the major rationale for performing assessment.

References