A history of disputes about reading instruction

Published: Wednesday, 06 November 2013

Dr Kerry Hempenstall, Senior Industry Fellow, School of Education, RMIT University, Melbourne, Australia.

Over the past thirty years, there has been considerable controversy over the competing emphases to beginning reading known as Whole Language and phonics. To provide a context for the debate, it may be helpful to consider the history of disputes about reading instruction, particularly as the instructional methods may differentially affect at-risk students. This article commences with a brief discussion of the advantages and challenges of our English alphabetic writing system, and of the literacy issues associated with it. Identification of the major attempts to deal with the complexity of our writing system is followed by a history of the search for the most efficacious means of evincing reading development. An examination of early research efforts, such as 'The Great Debate', 'The USOE Study', 'Follow Through', and 'Becoming a Nation of Readers' helps illuminate the current debate by highlighting which issues are from the past but as yet unresolved, and which are novel. A thread throughout the paper involves the role of educational research in influencing practice in beginning reading instruction.

Public interest surrounding the extent of literacy failure is at a high level currently, and in Australia, is reflected in the Federal Government's decision to allocate additional federal funding ($6.5 billion per year) to fund reforms in teacher education and providing further resources to address the needs of the disadvantaged (Topsfield, 2013). These reforms have been partly driven by the unsatisfactory results of the relatively recently introduced national testing (National Assessment Program – Literacy and Numeracy), and by similarly concerning results in the Progress In Reading Literacy Study. This was the first occasion on which Australia participated in this international assessment.

The tests revealed almost 25 percent of Year 4 children in Australia failed to meet the standard in reading for their age, to the shock of many educators and governments. In an open letter to federal, state and territory education ministers and their opposition counterparts, a group of 36 educators, scientists and clinicians call for a ``vast shake-up at all levels of teacher training'' to ensure children are taught to read properly (Ferrari, 2012).

The findings from such ventures, along with earlier studies by agencies such as the Australian Council for Educational Research and a report entitled National Inquiry into the Teaching of Reading (Department of Education, Science, and Training, 2005) have evoked sufficient concern within the community and among some education stakeholders to produce political action at the federal level.

So, community interest may be high at present; however, literacy and the role of schools in promoting it has had a fiery history in the educational community for almost two hundred years. Unfortunately, there has never been a consensus within the education community concerning the existence, definition, or extent of a literacy problem, and on appropriate methods of solving any such problem. This lack of unity has produced only fragmented intervention efforts, precluding the focused approach necessary to effectively address the systemic dilemma of illiteracy. The broad scale assessment in recent years, whilst remaining controversial, has produced illiteracy incidence figures that make it difficult to remain comfortable with the view that our system needs no improvement. The implications of such a failure rate threaten the welfare of the individuals involved and the nation as a whole.

Similar findings have galvanised action in both Great Britain (Department for Education and Employment, 1998; 2002; Department of Education and Skills, 2006) and the USA (National Assessment of Educational Progress, 2013; National Reading Panel, 2000).

A major continuing focus for dispute involves the significance of phonic strategies in beginning reading, and as an overarching theme, the role of educational research in influencing educational policy and practice. In each of these countries, federal policies have begun to reflect the converging results of exhaustive research analyses. These analyses have observed that phonics has been relegated to a minor role in beginning reading instruction. Further, their analyses have led to the conclusion that phonics should take a more central role in beginning reading, and they have instituted national policies to promote its occurrence (with varying degrees of success).

An examination of the history of reading disputes may be useful as it places the current debate in a broader context, and indicates how some contemporary issues are similar, or analogous, to those of earlier periods.

As far as we know, spoken communication has existed for as long as our species has developed relationships. How languages began is unknown – perhaps initially from imitation of the sounds heard in the natural environment, followed by invention of other sounds to encompass the many additional requirements of communication among the members of an intelligent species.

There are now at least 3000 different languages spoken in the world, yet the sounds that humans can produce are quite limited, and thus most languages require fewer than 50 distinct phonemes (Davis, 1988). The observation that all humans have developed language systems has prompted the suggestion that speech is biologically driven, and some (Smith, 1978) have extended this proposition to include written language as an analogous and equally natural process. This latter argument has always been highly contentious, and has profound implications for the design of reading instruction.

The model also assumes that reading (and writing) are natural parts of the same language process which enables the development of speech. Learning to read and write would be just as effortless and universal if the tasks were made as meaningful as is learning to talk. While the vast majority of children learn to speak with reasonable facility, a sizeable proportion of children do not learn to read well. One can recognize the principle of naturally unfolding development in Goodman's (1986) explanation for the disparity in ease of acquisition between speaking and reading. According to Goodman, reading problems occur through the breaking down of what is naturally a wholistic process into subskills to be learned and synthesized. Ironically, given all that is written today about “instructional casualties” (i.e., educational problems caused primarily by inadequate instruction (Reid Lyon, 2005; Wendorf, 2004) that includes whole language teaching), Goodman attributed reading failure to phonics teaching.

Liberman and Liberman (1990) dissected the equivalence assertion, and provided a strong counterargument. The Rousseauian perspective of naturally unfolding development involved the assumption that children have an innate developmental script that leads them (though perhaps at differing rates) to competence (Weir, 1990). Rousseau believed that unfettered maturation would allow the child to develop knowledge unaided. There is a readily apparent similarity in the unjustified Whole Language assumption that learning to read is essentially equivalent to learning to speak, and will develop as effortlessly as did speech. His argument that society should not interfere in the natural development of children generally, was paralleled by his view of the role of education. “Give your pupil no lesson in words, he must learn from his experience” (Rousseau, 1964 cited in Weir, 1990, p. 28). The Whole Language philosophy that assigns to the teacher the role of concerned facilitator, and that decries teacher directed instruction as harmful or unproductive can be readily sourced to this Rousseauian view.

Whole Language: What is it?

Goodman (1986) describes Whole Language as a philosophy rather than as a series of prescribed activities. Thus, Whole Language teaching consists of those activities a teacher with a thorough understanding of the philosophy would use. The teacher aims to provide a proper environment that will encourage children to develop their skills at their own developmentally appropriate pace.

This makes it difficult to describe what actually occurs in a Whole Language classroom, or whether there is any consistency from classroom to classroom that would enable an observer (other than one imbued with the philosophy) to recognize that the approach was indeed Whole Language. This reticence about detail is evident in numerous journal articles (Cambourne & Turbill, 2007; Smith, 1991; Newman, 1991; Johnson & Stone, 1991). There is a strong emphasis on principles, for example, the benefits of a natural learning environment (Goodman, 1986), and of exposure to a literate environment (Sykes, 1991). McGee and Lomax (1990) include authenticity, risk taking, choice, and empowerment as central concepts. Mills and Clyde (1990, cited in Johnson & Stone, 1991) provide a similarly difficult-to-operationalise outline of the Whole Language philosophy as evidenced in classrooms.

Highlight authentic speech and literacy events; provide choices for learners; communicate a sense of trust in the learners; empower all participants as teachers and learners; encourage risk taking; promote collaboration in developing the curriculum; be multimodal in nature; capitalise on the social nature of learning; encourage reflection. (p. 103)

What is evident from reading Whole Language literature is a perspective that children's educational progress is inner-directed, rather than environmentally evoked. Hence, its protagonists' aversion to formal instruction is understandable, though blinkered when the empirical research on effective teaching is considered.

Amongst educational philosophies, Whole Language is a relatively recent addition, and is derivative rather than unique. It has similarities to meaning-centred approaches in the manner in which it addresses reading - in that both support the primacy of meaning in the reading process. There are also similarities to the whole-word approach in the antipathy towards teaching sub-word analysis skills. However, whole language is rather more complex than simply a whole-word approach.

Writing Systems: From Logograms to the Alphabet

Attempts throughout history and across cultures to communicate in a visual format have varied in the style that the messages take, and to a greater or lesser extent these visual systems have met the written communication needs of the time. Apart from the requirement of communicating the author's intent, the system should be capable of expressing the full range of ideas, emotions, and actions for that culture. Furthermore, any system, if it is to be available to the general population (not only for an elite), should be easy to write/draw, and have a manageable number of symbols. It is not surprising then that, over time and over different cultures, numerous different systems have evolved. The earliest attempts involved pictures (e.g., cave drawings), and at their simplest they were quite effective when the writer had at least rudimentary skills. Complex ideas however were more difficult to draw skilfully and portray unambiguously; in fact, many ideas cannot be portrayed by drawing, for example, democracy.

Agreed-upon symbols evolved to overcome this problem, at least within geographical regions, but did not have the universal comprehensibility of, for example, a drawing of a horse. Such symbols (called logograms) are slow to reproduce and, as each is unique, require impressive memory capacity. The Chinese have at least 40,000 logograms (Rayner & Pollatsek, 1989), far more than any scholar could manage.

There is some evidence that the skills demanded of readers differ depending on the writing system in use. Huang and Handley (1994) noted that learning to read Chinese required less phonological awareness than was necessary in an alphabetic system such as English (and correspondingly greater visual skills). Such findings have important implications for the most appropriate instructional emphasis in initial reading. It is suggestive of the need to ensure the development of phonological awareness in students embarking upon beginning reading in an alphabetic system. In fact, there is now an irresistible body of evidence pinpointing phonological skills as powerfully causal in reading development. An examination of this evidence is beyond the scope of this paper; however, there are numerous reviews (Ball 1993; Hempenstall, 1997; Smith, Simmons, and Kame'enui, 1995).

About 4000 years ago the interest in word look-alikes shifted to word sound-alikes. Thus, rather than symbols representing words-in-picture, they could represent the sounds-in-words, initially through the use of syllables.

This had obvious advantages in economy because the same syllables appear in many words, and because all words are readily decomposable into syllables. The Korean written language system is partly a syllabary, containing several thousand syllables, and English contains about five thousand (Adams, 1990). An emphasis on syllables, however, means that each reader and writer must learn and memorise a large number of syllables that have no pictorial meaning. This requirement would restrict written communication to an elite. To avoid this restriction, the Korean written language (a fifteenth century invention) also incorporates an alphabetic system.

As the evolution of written systems continued, the requirement of general accessibility of the written system led to the association of one symbol, or letter, with each basic speech-sound or phoneme. Just as syllabaries reduced the memory load required by logograms, so the use of letters made possible reproducing any word in far fewer symbols than were required by syllabaries. This made the task of learning to read and write far more accessible to the general population, and the alphabet was recognised as one of the more significant of human inventions.

Problems of Written English

Written communication was developed mainly as a means of making inventories of ordnance and cargo, and was available to only a select few. Over time, it became a means of timeless communication of the prosaic and the profound, allowing almost any person to master its techniques.

Unfortunately this expectation has proved rather difficult to fulfill. A problem for an alphabetic system like English is the lack of a one-to-one correspondence between letters (graphemes) and sounds (phonemes) (Adams, 1990; Rayner & Pollatsek, 1989). This is especially evident with vowels – we have more than a dozen sounds represented by only five letters.

Our oral language has changed markedly from Old English (which was quite regular). Old English was Germanic in origin (Francis, 1965), but new words and sounds have entered our language, mainly from Latin, Greek, Nordic, and French. These new sounds and sound-combinations have to be encompassed within a print system that is unchanging – thus leading to the irregularities that are the bane of young (and not so young) readers and writers. The rules for letter-sound correspondence do not always provide the means to accurate decoding. For example, English has 1,120 ways of writing our 40+ sounds. In Italian, there are 25 sounds written in only 33 combinations of letters (Paulesu et al., 2001).

Perhaps the opposition to phonic instruction would not have been so great had the English written language been more regular. This is certainly the case in countries whose written language is orthographically transparent, such as Spanish, Italian and Finnish (Gonzalez & Valle, 2000), in which the teaching of phonics is considered essential.

These irregularities have led to several unsuccessful attempts at reforming the alphabet, the most famous being George Bernard Shaw's attempt in the nineteenth century, and the introduction of ITA (Initial Teaching Alphabet) in the twentieth century. These attempts are based on the principle that one-to-one correspondence between letter and sound will make phonics instruction more effective, and learning to read easier. Downing (1979, cited in Adams, 1990) reported on a large scale British study that found this to be true for ITA – students were more readily able to develop understanding of the alphabetic principle when taught using the ITA method of reading instruction. The counter-argument to such reform is that, while the ITA's phonemically regularised alphabet may aid beginning readers, it would be at the expense of skilled readers who are able to gain important and meaningful information from the traditional orthography. Under ITA, for example, homophones would have the same spellings, making comprehension more difficult; and intra-word conditional redundancies (an element in skilled word recognition) would be unavailable to the reader. Garner (1962, cited in Gibson & Levin, 1978) argued that letters are more constrained (and thus

more predictable) than words. These conditional rules about clusters of consonants, and the allowable number of vowels in a sequence reduce uncertainty; thereby, facilitating word recognition. It is the automatic simultaneous activation of intra-word units that distinguishes skilled readers (Roth & Beck, 1987). For example, in a word that begins T(consonant)(vowel) there is a very strong probability that the consonant is 'h'. Thus the effect on existing readers of such a reformed orthography may be to decrease reading speed and comprehension for the majority, at least in the short term. Whether such reform would be advantageous overall is irrelevant given that the disadvantages would fall on adults (the decision makers) who would be required to relearn the reading process.

An alternative strategy is to teach beginning readers the new orthography, and then teach the traditional form as they master the principles of reading. There is some doubt (Crowder & Wagner, 1992; Groff, 1990) whether such reform would be worth the trouble, as longitudinal studies (Rayner & Pollatsek, 1989) that have compared students in traditional and reformed systems found no significant between-group differences by the time that the transfer back to traditional orthography was completed.

Perhaps the most enduring outcome of the bold attempt at reform will be the recognition that an early emphasis on learning the alphabetic principle is most efficacious in beginning reading instruction (Chall, 1967).

We are left with an English system of 26 letters, and about 44 phonemes that can be spelled in hundreds of ways (Pollack & Pickarz, 1963). The permutations this allows makes learning to read and teaching beginning readers formidable tasks. Confusion arises from words that look alike but are sounded differently (tough, bough, cough, dough), and words that look different (mail, male) but sound alike.

Letter confusions are also common in beginning readers (Rayner & Pollatsek, 1989). It is probably the first time that they have encountered an object the name of which changes when it is rotated. A chair remains a chair when it is rotated, but a 'b' becomes a 'd' when rotated about the vertical axis, and a 'p' when rotated about the horizontal axis. Letters may be upper-case or lower-case, and in differing script forms. To further complicate the issue, the sounds of individual phonemes are not precisely maintained when they are used in words, they are influenced by the letters around them. Thus the phoneme /d/ has a different sound when followed by /oo/ than when followed by /i/. The converse is also true – there are not three distinct sounds in 'cat'. Only by a learned process of conscious analysis can one detect these individual phonemes (Stanovich, 1993), and it is a skill that eludes about 30 per cent of first grade readers (Adams, 1990). Phonemes in spoken words are really cognitive concepts rather than discrete sonic entities.

Despite the difficulties imposed by our system of writing, most children do learn how to read with at least reasonable proficiency. This is true over time, and across different languages, and systems of teaching; however, we are beginning to understand that not all systems lead to equivalent outcomes. The role of teaching, then, is to provide the opportunity, the encouragement, the environment, and the instruction appropriate to beginning readers needs so that learning occurs optimally. The approaches adopted by educators have been many and varied, but a major focus has been the degree to which strategies involving intra-word analysis are necessary in the development of reading.

The Problems of Literacy

That literacy is highly valued in a democratic technological society is readily apparent. From inquiries and policies at various levels of government, from media interest, from employer – expressed concerns, through to parental involvement - it is evident that the goal of literacy for all is of considerable importance in our society. Recognising what constitutes literacy is rather more difficult, however.

Literacy can be considered as set of skills rather than a unitary concept (Stedman & Kaestle, 1987), and people vary in the number of skills, and the degree of their mastery of those skills. Stedman and Kaestle distinguish between reading achievement and functional literacy. They define reading achievement as those skills that are taught and assessed in schools – from learning to read words already in their lexicon through to complex critical and interpretive skills. The overlapping dimension, functional literacy, implies the ability

to comprehend written communication outside school – in work, recreation and general societal participation. Deficits in functional literacy may be evident in an inability to read product labels, recipes, medications, traffic signs, street names, transport schedules, children's homework, school reports, and emergency phone numbers.

Kirsch and Guthrie (1984) suggest that the demands of the workplace for literacy skills is increasing, and jobs without a strong literacy requirement are becoming rare.

Similarly in A Nation at Risk (National Commission on Excellence in Education, 1983) it was argued that literacy demands are outstripping supply. It is more apparent than ever before that our society requires an educated population; yet, our rate of early school leaving has not reduced over the last decade, and under-developed literacy and numeracy skills are strongly implicated in early school leaving (Business Council of Australia, 2003).

Governments recognise that unless the nation's workforce is skilled, the economy will always be under threat because of intense global competition. Parents are aware that their children's economic and social futures depend on the capacity to respond productively to a rapidly changing workplace. We live in a complex world in which no country can afford to be insular. Nation's policies have an impact at a global level in addition to their domestic impact. Citizens' awareness of national and international issues is crucial to the development of a nation's policies; only an educated population can provide the needed support and guidance for their political leaders in a democratic society.

An obvious question is whether schools, as they are currently structured, are capable of meeting the increasing literacy requirements of our society. In addition to the increased requirements for literacy, schools are pressured to include many tasks in their curricula that formerly were considered family responsibilities, for example, health issues including sex education, drugs, smoking, fitness, skin cancer, and youth suicide. A focus on these questions concerning tasks and resources in education, rather than acrimonious debate on whether standards are declining, encourages a more forward-looking, constructive approach to achieving improved student literacy through a mutual emphasis on teaching/learning processes.

The Teaching of Reading: Early Controversy

The Emergence of Meaning-Centred Approaches

The first teachers of reading in English were priests in the seventh century. Children were taught the alphabet, syllables, and the Primer, or Prayerbook (Davis, 1973).

Most reading was religious, and the ability to read was restricted to relatively few. With the invention of the printing press in the sixteenth century the written word became much more prevalent, although the Bible was the only book available in most homes. Thus, reading was first promoted by religious authorities as a means to one end (salvation), and only later was considered important by governments, as a means to a quite different, secular end – an educated, democratic society.

The phonic technique of teaching component skills, and then combining those skills was the norm until the mid-nineteenth century (Adams, 1990). It followed a sequence of teaching upper-case and lower-case letter names, two-letter and three-letter combinations, monosyllabic words, multi-syllabic words, phrases, sentences, and finally, stories. Phonics is an approach to teaching reading that aims to sensitise children to the relationships of the spelling patterns of our written language to the sound patterns of our oral language. It is not a single method, however, as decisions need to be made regarding the timing of its introduction, the method of delivery, whether explicitly taught, or simply implied, taught in isolation, or solely in the context of literature, how many, and which, rules are appropriate.

It was not until 1828 that Samuel Worcester produced a primer that borrowed a European idea of teaching children to recognise whole words without sounding them out.

It is not very important, perhaps, that a child should know the letters before it (sic) begins to read. It may learn first to read words by seeing them, hearing them pronounced, and having their meanings illustrated;

and afterward it may learn to analyse them or name the letters of which they are composed. (Crowder & Wagner, 1992, p. 204).

Support for this view came from James Cattell in 1885 in his assertion that whole word reading was more economical (Davis, 1988); and later, from the Gestaltists who considered that the overall shape of the word (rather than the summation of the sound-parts) should provide the pre-eminent clue for young readers. An assumption behind this approach was that beginning readers should be taught to read in the way skilled readers were thought to do. Given the belief that skilled readers associated meaning directly onto the orthographic image, it follows that there would be time saved by showing beginners how this was achieved. The opposing view was that reading should be viewed as a developmental process in which the early stages of developing the alphabetic principle are necessary for later skilled-reading, even though those early skills may be rarely needed at the later stages. A further assumption of what became known as the whole-word approach was that the knowledge of letter-sounds would naturally follow once whole-word recognition was established (Smith, 1978). It was not until some time later that doubt began to be expressed about the effects on some children of this whole-word initial emphasis.

Unfortunately for such children, the consequence of the primacy of the whole-word method was an inability to decode unfamiliar words (Tunmer & Hoover, 1993).

The major reason for the length of time that elapsed before empirical judgements could be made about the relative merits of the contrasting teaching emphases relates to the dearth of investigators engaged in such research until comparatively recently, and to the abysmal quality of most educational research (Feuer, Towne, & Shavelson, 2002).

The History of Educational Research into Teaching Reading

It was not until the first two decades of this century that educational research began in earnest. The development of formal reading tests, and the recognition that education was a fertile ground for research, led to many investigations into such topics as remedial approaches, individual differences in development, test development, silent reading vs oral reading, and reading-readiness.

Although this research was in its infancy, early findings (often unsubstantiated by other research) were quickly adopted by book publishers keen to exploit the new markets that mass education provided. A number of texts based on whole-word teaching were published, and the method became very influential throughout the 1930s and 1940s. It appeared to offer a curriculum sensitive to the developmental needs of students, and one that would be both more attractive to teachers than phonics drill and more interesting to the rapidly increasing numbers of students attending school beyond the primary level.

The whole-word model, as outlined by Chall (1967), involved introducing words through their meaning.

Words should be recognised by sight, using the cue of their shape and length. A fall-back strategy relied on deducing meaning from other clues such as pictures, or from the context. Phonic strategies were considered potentially harmful, to be used only as a last resort. Even then, phonics was usually employed to provide only partial cues, such as from attention to a word's first or last letters.

Systematic teaching of phonic strategies was antithetical to the wholistic nature of such meaning-oriented approaches.

Because teaching should not take as the unit of instruction anything other than meaningful text, any substantive phonic skills developed by students would necessarily be self-induced and idiosyncratic.

The approach was taken even further when the wholesentence, and then the whole-story became the units of study. In the sentence method, the child looked at the sentence being read by the teacher. This reading was followed by a focus on particular words in that sentence.

In the whole-story method, the story was read to the child by the teacher before sentences and words were addressed.

These emphases were designed to make pre-eminent the meaning of print rather than its construction, and was thought to be more interesting for the child, thereby enhancing learning.

Unfortunately, as the unit of analysis enlarged the more necessary it became for students to rely on memory.

Some texts began to use controlled vocabulary in the early reading stages, but the problem of decoding unfamiliar words was merely postponed. It was anticipated that a self-directed attention to word similarities would provide a generative strategy. However, such expectations were all too frequently dashed, as for many at-risk children progress came to an abrupt halt around Year 3 or 4 when an overwhelming number of unfamiliar (in written form) words are rapidly introduced. This apparent stalling of progress became known as the fourth grade slump (Chall & Jacobs, 1983; Hirsch, 2003). The number of words a child requires to comprehend grade level text in Year 2 was estimated by Carnine (1982) as between three and four hundred, and in Years 3 and 4 between three and four thousand. Share (1995) estimates that the average fith year student encounters about ten thousand new words.

Strategies that rely upon memory-for-shapes of words, or picture-clues, or context-clues become unproductive (Spear-Swerling & Sternberg, 1994). Depending largely on their visual recognition store of word shapes, students too often do not develop any generative strategy for the decoding of novel words. It is true that many children develop a working understanding of the alphabetic principle despite the absence of explicit instruction; however, those students who did not have the 'Aha!' experience tended to be left floundering without the structure necessary to progress (National Reading Panel, 2000).

Prior to the whole-word dominant period, it was oralreading that was most commonly taught and tested; however, with the increased emphasis on reading for meaning, silent reading began to increase in popularity.

Unfortunately the cost of abandoning oral reading was the loss of information available to the teacher about progress and problem areas. This allowed reading errors to be practised to the point of being firmly established.

In addition, oral reading assists readers to become more familiar with those words whose spellings do not match their pronunciations (Adams, 1990) – it assists students to become aware that written language provides the same opportunities for communication as does its oral form; and, in beginning readers, it leads to higher word recognition and comprehension scores (Carnine & Silbert, 1979). For older, more skilled students, the primary mode of reading is silent.

The seemingly obvious solution involves a suitable balance so that both oral and silent reading opportunities are regularly scheduled at the appropriate reading stages. However, to some theorists, oral reading does not provide an authentic experience, because meaning may be compromised. "The basic mode of reading is silent.

Silent reading does not place constraints on the reader" (Barmby, Bonham, Lawry, & Nisner, 1986, p. 35). Even today, many schools schedule daily silent reading under a variety of acronyms while considering unnecessary the provision of opportunities for corrective feedback through regular oral reading. The presumption that practice makes perfect – that increasingly skilful reading will occur as long as the child engages in reading regularly is misplaced.

The National Reading Panel (2000) cautioned that there was insufficient evidence that silent reading plays a part in reading development. In the absence of feedback, practice is likely to make errors permanent (Fields & Kempe, 1992), and this is especially true for at-risk students (Kame'enui & Simmons, 1990).

The debate over code-emphasis vs meaning-emphasis has always been vitriolic. During the 1840s, the Secretary of the Massachusetts Board of Education reporting on code-based (phonics) teaching colourfully

described “the odor and fungousness of spelling-book paper” from which “a soporific effluvium seems to emanate ... steeping (the child’s) faculties in lethargy”. By contrast, meaning-oriented lessons “will be like an excursion to the fields of elysium compared with the old method of plunging children, day by day, for months together in the cold waters of oblivion, and compelling them to say falsely, that they love the chill and torpor of immersion”. (Adams, 1990, p. 22) In the 1950s, the first real challenge to the whole word approach was initiated by Rudolph Flesch, who, in an emotion-charged attack, wrote: It seems to me a plain fact that the word method consists essentially of treating children as if they were dogs. It is not a method of teaching at all; it is clearly a method of animal training. It’s the most inhuman, mean, stupid way of foisting something on a child’s mind. (Flesch, 1955, p. 126) Flesch’s arguments were fuelled by aligning them with a perceived threat to democracy posed by an alleged decline in reading standards in the working class. He hinted at conspiracies to disempower sections of the community by deliberately using methods of teaching that were ineffective. “The American dream is, essentially, equal opportunity through free education for all. This dream is beginning to vanish in a country where public schools are falling down on the job” (Flesch, 1955, p. 132). Flesch’s call for a return to phonics teaching had an enormous impact – the book was a best seller, and perhaps for the first time, parents began to express publicly a desire to be involved in educational decision-making.

Community interest has continued from that time up to the present, and some similarities can be seen in the current phonics and Whole Language debate. However, this was the first real taste of public accountability in education, and it had a significant impact on researchers, publishers and politicians. Courses on reading became more important in teacher training, research intensified, and government inquiries into literacy became regular events. Publishers began producing a wider variety of reading programs. They varied from code-emphasis to meaning-emphasis, and various combinations of these features for those wishing to take an eclectic stance on the issue. There is an analogous situation currently as some have argued for a “balanced literacy” (Moats, 2000) that is intended to incorporates the best elements from both the phonics and Whole Language approaches.

Flesch’s arguments had moved beyond the pedagogical to the political. The linking of phonics with democracy was a most effective strategy, though not one that endeared him to educational historians. His actions, however, were based on his assessment of which approach was more effective in teaching children to read. The writings of a number of the leaders of the Whole Language movement (Edelsky, 1990; Goodman, 1989) display a similar interest in politics; indeed, politics appears to be the primary concern of these writers rather than effective teaching practice (as Flesch had intended his political stance to evoke). Questions of teacher or instructional effectiveness are less important to such advocates than are the objectives of personal liberation for students and for society. “Whole Language ... has human emancipation as its goal” (Shannon, 1994, p. 99). Given the conceptual disparity between these major objectives it is unsurprising that genuine dialogue between Whole Language advocates and those seeking instructional sophistication was never achieved.

In Flesch’s time, there was still little systematic evidence about the relative effectiveness of the two major emphases across the broad population, and much debate centred on philosophical issues. R. H. Thouless might have had just such an issue in mind when he formulated his Law of Certainty. It can be summarised by the observation that when there is cause for doubt about a particular belief, or conflicts between approaches that are not readily resolved, one may reasonably expect that most people would adopt a position of caution. However, Thouless observed that such uncertainty seemed to strongly polarise people’s views so that more are likely to hold extreme views than to hold a moderate position. Thus, supporters may clutch even more strongly to a belief about which there is doubt, while detractors focus strongly on the apparent negative aspects of the belief, disregarding any positives. This observation may partly explain why educational policy-making has been subject to such extreme pendulum swings.

The Great Debate

During the 1960s, Jean Chall (1967) was an important figure because she accepted the Herculean task of developing a scientific study to test the effectiveness of various approaches to reading. The outcome of her

work, Learning to read: The great debate was published in 1967, and her conclusions were then, and remain now, controversial.

Having analysed twenty basal level reading programs across 300 classrooms in three countries, and having studied the literature (such as it was) on effectiveness comparisons of phonics and whole-word approaches, she concluded that systematic teaching of phonics tended to produce better word recognition, spelling, vocabulary and comprehension in all children, not only those from the at-risk groups (such as students of lesser intelligence, or those from lower socioeconomic backgrounds). Chall's detractors (e.g., Carbo, 1988) have disputed her conclusions, arguing that much of the research she reviewed had a variety of methodological flaws involving non-standardised tests, non-random selection, and inadequate program descriptions. Despite the criticisms, Chall's contribution was influential in affecting the weight of opinion regarding phonics (at least among researchers and some empirically-minded educators), and in stimulating a great deal of subsequent research.

Some of this later research continued to be criticised as flawed but, in general, improved rigour began to be a feature of the design of educational studies. The results of her research were, however, less successful in altering the products of the publishers of beginning reading texts, and in influencing education bodies to promote practices of proven effectiveness in their domain. The failure of research-based knowledge to have an impact upon educational decision-makers continues to be lamented (Carnine, 1995, 2000, 2003; Hempenstall, 1996; Stone, 1996).

The USOE Study

In the USA, the strength of public interest ensured that concern and research funding from governments was forthcoming. Large scale projects followed throughout the late 1960s and 1970s. The US Office of Education Co-operative Research Program in First Grade Reading Instruction was designed to overcome the criticisms of Chall's work, and to extend the research questions.

Which approaches to beginning reading work best? Does reading-readiness affect program effectiveness? What characteristics of communities, schools, teachers and students are correlated with better outcomes? (Bond & Dykstra, 1967). Twenty seven separate projects involving hundreds of classrooms were established – each informing on an element of the research questions in carefully designed studies coordinated by Bond and Dykstra. The findings echoed those of Chall regarding the importance of phonics teaching, but also argued that some balance between phonics and meaning-emphasis was most productive.

Well-designed programs were found to be effective for students despite their differing degrees of readiness. This finding was important because a whole industry of reading-readiness training was springing up. It was based on the assumption that children should not be taught reading until they had mastered a variety of visuo-spatial, language and motor skills. However, this underlying process approach represented an educational cul de sac. There were intractable problems in ascertaining precisely what these core fundamental skills were, accurately assessing them, and teaching them effectively.

Unfortunately, though some children became adept at drawing lines accurately within parallel boundaries, there was no reliable impact on reading progress (Arter & Jenkins, 1979; Hammill, 2004). In fact the meta-analysis performed by Kavale and Forness (1985) produced an overall effect size for perceptual-motor training of 0.08, which is considered a small effect (Cohen, 1988). What was not readily apparent at that time was that learning to read was the most effective way to master many of those skills – hence valuable instructional time was better spent on the target task. "If the goal is for children to learn a particular skill (like reading), it is more efficient to teach it directly than to expect it to transfer from other learning" (Singer & Balow, 1981, p. 107). Kavale's (1990) summary of research into direct instruction concluded that they are five to ten times more effective for struggling students than are practices aimed at altering unobservable learning processes such as perception.

In an analogous sense, the reading-readiness debate that gave primacy to the students' developmental stage in the ascription of when and what to teach, is being mirrored today in the interest among some

developmentalists in socalled “learning styles” (Carbo, 1992; Dunn, Beaudry, & Klavas, 1989). In this view there are important differences among children in their processing skills that require the tailoring of instruction to take account of them. If we accept the dubious proposition that learning styles are genuine and important variables in learning, and further, that they can be validly and reliably measured, then matching the instruction to the individual preference should produce superior learning. The approach has considerable intuitive appeal, and is the subject of a great deal of interest in education. As regards reading however, there is little evidence that such matching enhances the process of learning to read (Snider, 1992; Stahl & Kuhn, 1995). A meta-analysis by Kavale and Forness (1987) produced an overall effect size for modality instruction of 0.14, which is considered a small effect (Cohen, 1988).

It was significant that, in the Bond and Dykstra 1967 study, the meaning-oriented approach (out of which evolved ‘language-experience’ and ‘Whole Language’) did as well as basal (without phonic-emphasis) programs with high-readiness students but less well with low-readiness students. The adverse finding has been echoed over the past twenty years as ‘Whole Language’ approaches are frequently criticised because of their apparent ineffectiveness with at-risk students (Bateman, 1991; Gersten & Dimino, 1993; Stahl, 1990; Stahl & Kuhn, 1995; Vellutino, 1991; Yates, 1988).

Bond and Dykstra’s findings were concise regarding the characteristics of communities, classrooms, teachers and students that were predictive of successful reading instruction. The major student predictor was not intelligence but knowledge of letters (predating the now acknowledged predictive power of phonological skills).

The other finding that perhaps played a part in the rise of the “effective-teaching” movement (Rosenshine & Stevens, 1984) was the importance of the method of delivery (in addition to the issue of content). The conclusion that teacher variables have a significant influence on student success was very important at a time when teacher differences were considered by many to be of little significance.

“The implication is that to improve reading achievement we must improve both programs and classroom delivery. Each seems to contribute separately and significantly to children’s progress” (Adams, 1990, p. 43).

In the following largest educational study yet conducted, a model known as Direct Instruction successfully combined an explicit phonics emphasis with a teaching approach emphasising explicit, systematic instruction of the type described in the ‘effective teaching’ research.

Follow Through

The next major study was federally funded in the USA in the late 1960s. It arose because of a concern about the poor educational outcomes for disadvantaged students. Entitled ‘Follow Through’, it was aimed at the primary school stage, and was designed to determine which methods of teaching would be most effective for disadvantaged students throughout their primary school career. It followed an early-intervention project called ‘Head Start’ that had as its goal the preclusion of educational disadvantage. The results of ‘Head Start’ interventions unfortunately were not durable, and the model failed to achieve its ambitious objectives. Unfortunately, that continues to remain the case (Puma et al., 2012).

The impact of this expensive failure was felt by ‘Follow Through’. Though initially intended as a massive intervention, it was reduced in scope to that of a study to assess how best to maintain and build on ‘Head Starts’ fragile gains. It remained, however, a huge study – involving 75,000 children in 180 communities over the first three years of their school life. It continues to be the largest educational experiment ever undertaken, extending from 1967 to 1995, at a cost of almost a billion dollars. There were nine major competing sponsors covering a broad range of educational philosophies. They included child directed learning, individualised instruction, language experience, learning styles, self-esteem development, cognitive emphasis, parent-based teaching, direct instruction, and behavioural teaching. The models can be reduced to three distinct themes – those emphasising either basic academic outcomes, cognitive development, or affective development. The targeted basic skills included reading, language, spelling, writing, and maths.

The models that emphasised the systematic teaching of basic skills (Direct Instruction, and Behaviour Analysis), performed best. In reading, the Direct Instruction model, with its strong systematic synthetic phonics emphasis, had the most impressive results in both academic and affective areas.

There were criticisms that variability in implementation across sites made judgements of model superiority dubious, and that overall effects were too small to be pleased about (House, Glass, McLean, & Walker, 1978). Nevertheless when the data were re-analysed by several groups (House et al., 1978; Bereiter & Kurland, 1981; Meyer, Gersten, & Gutkin, 1983), the Direct Instruction (DI) model still produced the best gains. Later follow-up studies (Becker & Gersten, 1982; Gersten, Keating, & Becker, 1988) were completed over the following ten years, and added support to the argument that the superiority of the Direct Instruction model was real and significant.

To expect gains to endure over such a long period might be considered unrealistic, but Chall (1979) had argued that if children could master the decoding stage "... the knowledge and skills acquired are usually sufficient to become self-generative. That is, further growth can be achieved with practice on one's own" (p. 47). This concept was extended by Share (1995) when he described phonological recoding as a mechanism enabling self-teaching of the decoding of novel letter combinations.

Stanovich (1986) emphasised the role of practice by citing it as the major determinant of vocabulary growth after about Year 4, and even important in subsequent intellectual development. Thus the positive findings in the follow-up studies imply that early skill mastery had led to a continued interest and involvement in reading for those disadvantaged students who graduated from the Direct Instruction model. The DI model has been criticised (Schweinhart, Weikart, & Larner, 1986) for its strong emphasis on teacher-directed, scripted lessons, alleging a consequential over-reliance on teachers, and an inability to self-direct learning. However, follow up studies of the DI students showed "strong consistent long term benefits in reading" three, six, and nine years after students completed 'Follow Through' (Gersten, Keating, & Becker, 1988, p. 326). The effects were evident in higher achievement, fewer grade retentions, and more college acceptances than in comparison groups that had traditional education in the same communities. Oddly, the results of this huge comparison study had negligible impact upon educational policies.

Becoming a Nation of Readers

In 1985, Becoming a Nation of Readers (Anderson, Hiebert, Scott, & Wilkinson, 1985), a report of the Commission on Reading, examined the teaching of reading, reading problems, and likely solutions. The report endorsed approaches that included a strong, early, synthetic phonics emphasis.

The argument about the constituents of effective phonics teaching is currently being revisited, as there is now a developing acceptance of the importance of phonic strategies in beginning reading. While some Whole Language theorists still believe that any emphasis on phonics is unfruitful, or even harmful – "The rules of phonics are too complex... and too unreliable... to be useful." (Smith, 1992, p. 438), the major disagreement now revolves around the mode of teaching – not if phonics, but how phonics. Some acknowledge a role for phonics, albeit a secondary one. "Almost by definition, we can say that good readers are ones who use context efficiently, to reduce their reliance on visual cues and grapho-phonemic knowledge" (Weaver, 1988). Of those Whole Language advocates who see a role for phonics in a reading program, most argue that any word analysis skill development should occur only in the context of reading connected text (Weaver, 1988). See Iverson and Tunmer (1993) for a fuller discussion of this issue.

Thus the sort of systematic explicit phonics teaching envisaged by the report of the Commission on Reading is unlikely to be found in a modern Whole Language classroom.

"Empirical investigations may demonstrate, as miscue analysis has, that phonics is a distinctly inferior cuing system, one hardly deserving any privileged status (Strauss & Altwerger, 2007).

It is not that such teaching could not be included, but that it is proscribed by the major writers in that field (Edelsky, 1990; Goodman, 1986, 1989; Smith, 1992; Weaver, 1988). This antipathy toward systematic

synthetic phonics continues, and is not necessarily influenced by data. Dr Morag Stuart, a British reading researcher, in providing evidence to a parliamentary enquiry, noted: I nearly lost one school because the phonics-taught children were doing better than the non-phonics-taught children, and this head teacher said to me that she was ideologically opposed to taking part in a study which showed that phonics teaching worked. (Education and Skills Committee, 2004).

Henry (1993) argues that Whole Language's lack of explicitness regarding phonics militates against at-risk learners, as they are the least likely to develop their own phonic generalisations. A further problem for such students is that such unsystematic access to useful phonic principles leaves them without a firm basis for mastery, or with enough massed and spaced practice for incorporation to occur. It should be noted that all phonics are not equal; it is possible to employ phonics carefully and with parsimony; it is possible to do it ineffectively and excessively; and it is possible to do it in name only. A fuller discussion as to the important elements of such approaches may be found in Foorman, 1995; Groff, 1990; Hempenstall, 2002; Henry, 1993; National Reading Panel, 2000; Stahl, 1992.

The Becoming a Nation of Readers report urged that comprehension instruction also be explicit:

While questions during the preparation and discussion phases of a reading lesson are important, these do not substitute for active, direct instruction. In direct instruction, the teacher explains models, demonstrates, and illustrates reading skills and strategies that students might be using. There is evidence that direct instruction produces gains in reading achievement beyond those that are obtained with less direct means such as questions (Anderson, Hiebert, Scott, & Wilkinson, 1985, p.56)

Becoming a Nation of Readers defined as state-of-the-art a direct teaching model with an early phonics emphasis. It was critical of much of existing practice in beginning reading, in methods of teaching comprehension, and in a lack of systematic formal and informal assessment. In common with a number of recent commentators and researchers , this report did not consider that the early inclusion of phonics instruction precluded a parallel emphasis on meaning, or the use of authentic literature.

The Impact of Research on Practice

In the years following, researchers have begun to look more closely at specific elements of curriculum content, and methods of instruction to allow a more fine-grained analysis of what works best for whom, and at what stage.

Thus far the lack of impact of this research on educational practice has been of concern to many in the educational community (Carnine, 2000; Hempenstall, 1996; Stone, 1996). In earlier times, research findings were rarely conclusive, and it is understandable that such research results were not a major force in educational policy formulation. There is now a consensus among empiricist researchers about a number of issues crucial to reading instruction (National Reading Panel, 2000). Certainly if one examines empirically accepted findings such as summarised by Vellutino (1991), it is difficult to accept the status quo:

(a) The most basic skill in learning to read is word identification; (b) an adequate degree of fluency in word identification is a basic pre-requisite to successful reading comprehension; (c) word identification in skilled readers is a fast-acting, automatic, and in effect modular process that depends little on contextual information for its execution; (d) even skilled readers can accurately predict no more than one word out of four in sentence-contexts, indicating that the predictive role of context must be extremely limited; (e) because of limited facility in word identification, beginning and poor readers are much more dependent on context than are more advanced and good readers; (f) facility in alphabetic coding is critically important to the acquisition of skill in word identification; (g) phoneme awareness and facility in phoneme analysis are critically important to the acquisition of skill in alphabetic coding. Each of these generalisations is contrary to the approach to reading instruction currently advocated by Whole Language proponents (p. 442).

The idea that the results of empirical research should play a strong part in decision-making and policy development in any important area of human service is not novel; however, in the field of education, as in a

number of areas of human services, there has been relatively little weighting given to research (Gable & Warren, 1993). Rather, broad philosophical principles such as developmentalism (Stone, 1996) have been the determining factor in the direction taken, most evidently in the fields of teaching and learning. This preference for dogma over pragmatism was noted in a number of countries besides Australia (Stanovich, 1994), and led to the domination of the Whole Language philosophy in policies on teaching and learning in numerous countries.

The Australian House of Representatives Enquiry - The Literacy Challenge (1993) noted that Whole Language had Australia-wide support.

...virtually all curriculum guidelines on primary school literacy teaching produced are based on this approach... Virtually all teachers have undertaken the inservice training course, Early Literacy Inservice Course (ELIC), which is also based on a Whole Language approach to learning and literacy (p. 25)

Whole Language philosophy itself has been relatively impervious to the results of research. In fact, as McCaslin (1989) noted, Whole Language advocates assert that the research perspective itself is responsible for inappropriate teaching practices. Edelsky (1990) argued that "...procedural rigour in research design is no more than a thinly disguised demand that Whole Language be translated into terms that fit a skills model of reading and a positivist model of research" (p. 10).

This perspective is presumably responsible for the dearth of quantitative research by Whole Language advocates on the effectiveness of the approach (Klesius, Griffith, & Zielonka, 1991; Stahl, McKenna, & Pagnucco, 1994).

There are a number of researchers (Adams & Bruck, 1995; Chaney, 1990; Fields & Kempe, 1992; Gersten & Dimino, 1993; Heymsfeld, 1989; Mather, 1992; McCaslin, 1989; McGinitie, 1991; McKenna, Robinson, & Miller, (1990a, 1990b); Spiegel, 1992; Stahl & Miller, 1989; Stanovich, 1994) who consider that the investment in Whole Language has too great for the model to be completely displaced, and who seek a rapprochement, allowing Whole Language to take advantage of newer, demonstrably effective practices yet still retain its flavour.

However, Whole Language stalwarts view such gestures as reactionary, as they consider a skill development model incompatible with the essence of Whole Language (Edelsky, 1990; Heymsfeld, 1989; Goodman, 1989). This perspective continues to be evident in the writings of such Whole Language stalwarts as Cambourne and Turbill (2007).

Ball (1993) also notes the conflict between the Whole Language perspective and research. In her view the pedagogical battle between empiricists and Whole Language supporters is reflective of a broader debate evident in many of the social sciences. The major debate is between those who support a reductionist, positivist philosophy of science and those who rebel against that position – adopting a holistic, post-positivist, relativistic stance. To relativists, such as Weaver (1988), all empirical research is futile in determining teaching practice because the very performance of the research cannot avoid influencing the outcome, thereby confounding results.

Relativists view reality as phenomenological, that is, it has no existence independent of each person's unique individual perspective. They tend to favour ethnographic approaches such as case studies and classroom observation (McKenna, Stahl, & Reinking, 1994) as the appropriate means of enquiry. Empiricists view reality as "essentially cognitive transcending" (Rescher, 1982, as cited in Groff, 1990), and see ethnographic research as useful for raising ,rather than answering, questions about teaching practice.

Empiricism and Knowledge Claims

Those who support empirically-based models are likely to agree with Stanovich (1994) when he proposes that competing claims to knowledge (such as about models of teaching) should be evaluated according to three criteria.

First, findings should be published in refereed journals.

If research is to be useful it must be well designed, and able to justify its findings. When peer review is part of the process of research the well known taunt “research can prove anything you want” becomes less credible. Poorly designed studies are rejected (sometimes to appear in un-refereed journals).

Second, reported results should be replicated by independent researchers. A new model or strategy can be more sanguinely adopted when findings have been repeated in studies in which the researchers have no particular stake in the outcome.

Third, there is a consensus within the appropriate research community about the reliability and validity of the findings. This last criterion requires considerable reading across the field, but the frequency with which a particular study is cited in journal articles and accepted as legitimate provides one measure. Although the use of these criteria cannot guarantee infallibility, it does offer reasonable consumer protection against spurious claims to knowledge. While such a litmus test is appealing to empiricists, it is unlikely to impress those Whole Language proponents who view data with a different perspective, and make decisions on other grounds.

There is, then, a vast gulf separating empirical approaches to teaching and learning from those still favoured in many education systems (policy makers and education schools). The gulf transcends mere disagreement about effective strategies; the differences are at a more fundamental level and represent a significant challenge even to find a mutually acceptable framework to allow dialogue. There are significant changes notably through government pronouncements, but change at the level of the teacher has been difficult to institute, especially among committed Whole Language enthusiasts.

However, for most teachers the philosophical impasse may be less problematic. Vellutino (1991) believes that teachers are more pragmatic than are the major writers on Whole Language. He argues that teachers may espouse the current orthodoxy, while practising a hybridised approach – what Newman (1991) disparagingly described as “...teaching Whole Language in the afternoons” (p. 73). It is difficult to know how prevalent this is, although a number of researchers have noted the increased acceptability to teachers (and value to students) of including early phonemic awareness and phonic emphasis instruction (Castle, Riach, & Nicholson, 1994; Eldredge, 1991; Fields & Kempe, 1992). More recent research has only strengthened this perspective (Brady, Braze, & Fowler, 2011). Providing adequate training to teachers to enable them to make use of the research findings is another issue confronting educational reformers.

So, is Whole Language dead?

There is less flag waving these days, and terms such as balanced literacy instruction are more likely to be employed (Moats, 2000).

An outcome of this inclusiveness was a move away from using the term whole language and simply using literacy. "To us the principles underpinning the word literacy were similar but did not bring with it the negative connotations. … Whole language is still with us, strongly embedded in current curriculum, pedagogy and assessment strategies. Adversaries of whole language still complain that the term whole language may not be used however the philosophy is alive and well in each state system” (Cambourne & Turbill, 2007, p.23, 25).

In Australia, as is occurring in the US and Great Britain, dramatic curriculum change is occurring without the warring factions ever uniting. Governments have demonstrated an increasing interest in the establishment of national testing programs. In addition, governments have shown a distinct preparedness to examine the effectiveness of programs that compete for the scarce education dollar. This market perspective considers that student outcomes are determined more by education rather than by social factors. The oft-expressed view that students’ achievements will not be advanced unless poverty and disadvantage are first eliminated is unlikely to fall on receptive ears in future.

References

Adams, M.J. (1990). *Beginning to read: Thinking & learning about print*. Cambridge, MA: MIT Press.

Adams, M.J., & Bruck, M. (1995). Word recognition: The interface of educational policies and scientific research. *Reading and Writing: An Interdisciplinary Journal, 5*, 113-139.

Anderson, R.C., Hiebert, E. H., Scott, J.A., & Wilkinson, I. A. G. (1985). *Becoming a Nation of Readers: The Report of the Commission on Reading*. Champaign, IL: The Centre for the Study of Reading.

Arter, J.A., & Jenkins, J.R.V. (1979). Differential diagnosis-prescriptive teaching: A critical appraisal. *Review of Educational Research, 49*, 517-555.

Ball, E.W. (1993). Phonological awareness: What's important and to whom? *Reading & Writing: An Interdisciplinary Journal, 5*, 141-159.

Barmby, S., Bonham, D., Lawry, F., & Nissner, H. (1986). *Real literacy: Shifting the focus*. Vic, Australia: Maroondah Region Publication.

Bateman, B. (1991). Teaching word recognition to slow learning children. *Reading, Writing and Learning Disabilities, 7,* 1-16.

Becker, W.C., & Gersten, R. (1982). A follow-up of Follow Through: The later effects of the Direct Instruction model on children in fifth & sixth grades. *American Educational Research Journal, 19*, 75-92.

Bereiter, C., & Kurland, M. (1981). A constructive look at Follow Through results. *Interchange, 12*, 1-22.

Bond, G.L., & Dykstra, R. (1967). The co-operative research program in first grade reading instruction. *Reading Research Quarterly, 2*, 5-142.

Brady, S.A., Braze, D., & Fowler, C.A. (2011). Explaining individual differences in reading: Theory and evidence. New York: Psychology Press.

Business Council of Australia. (2003). *The cost of dropping out: The economic impact of early school leaving.* Retrieved from http://www.bca.com.au/upload/The_Cost_of_Dropping_Out.pdf

Cambourne, B. & Turbill, J. (2007). Looking back to look forward: Understanding the present by revisiting the past: An Australian perspective. *International Journal of Progressive Education, 3*(2), 8-29.

Carbo, M. (1988). Debunking the great phonics myth. *Phi Delta Kappan, 70*, 226-40.

Carbo, M. (1992). Giving unequal learners an equal chance: A reply to a biased critique of learning styles. *Remedial and Special Education, 13*(1), 19-29.

Carnine, D. (1982). *Approaches to reading: Psycholinguistics and Direct Instruction methods*. Paper presented at the Conference of the Australian Association for Direct Instruction, Kuringai CAE, NSW: Australian Association for Direct Instruction.

Carnine, D. (1995). Trustworthiness, useability, and accessibility of educational research. *Journal of Behavioral Education, 5*, 251-258.

Carnine, D. (2000). Why education experts resist effective practices (and what it would take to make education more like medicine). Washington, DC: Fordham Foundation. Retrieved from http://www.edexcellence.net/library/carnine.htm

Carnine, D. (2003, Mar 13). IDEA: *Focusing on improving results for children with disabilities*. Hearing before the Subcommittee on Education Reform Committee on Education and the Workforce United States House of Representatives. Retrieved from http://edworkforce.house.gov/hearings/108th/edr/idea031303/carnine.htm

Carnine, D., & Silbert, J. (1979). *Direct instruction reading*. Columbus, OH: Charles Merrill.

Castle, J. M., Riach, J., & Nicholson, T. (1994). Getting off to a better start in reading and spelling: The effects of phonemic awareness instruction within a whole language classroom. *Journal of Educational Psychology, 86*, 350-359.

Chall, J. S. (1967). *The great debate*. New York: McGraw Hill.

Chall, J. S. (1979). The great debate: Ten years later, with a modest proposal for reading stages. In L. B. Resnick & P. A. Weaver (Eds.), *Theory & Practice in Early Reading. Vol. 1*, (pp. 29-54). Hillsdale, NJ: Lawrence Erlbaum.

Chall, J. S., & Jacobs, V. A. (1983). Writing and reading in the elementary grades: Developmental trends among low-SES children. *Language Arts, 60*, 617-626.

Chaney, C. (1990). Evaluating the Whole Language approach to language arts: The pros and cons. *Language, Speech, and Hearing in Schools, 21*, 244-249.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.), Hillsdale, NJ: Lawrence Earlbaum.

Crowder, R., & Wagner, R. (1992). *The psychology of reading: An introduction*. New York: Oxford University Press.

Davis, A. (1988). A historical perspective. In J. Estill Alexander (Ed.) *Teaching reading* (3rd ed., pp. 532-553). New York: Scott, Foresman & Co.

Davis, W. (1973). *Teaching reading in early England*. London: Pitman.

Department for Education and Employment. (1998). *The National Literacy Strategy: Framework for Teaching*. London: Crown.

Department for Education and Employment. (2002). *The National Literacy Strategy: The first four years 1998-2002*. OFSTED Report HMI 555. London: Crown.

Department of Education and Skills. (2006). *Primary National Strategy: Primary framework for literacy and mathematics*. UK: Department of Education and Skills. Retrieved from http://www.standards.dfes.gov.uk/primaryframeworks/

Department of Education, Science, and Training. (2005). *Teaching Reading: Report and Recommendations*. National Inquiry into the Teaching of Reading. Australia: DEST.

Dunn, R., Beaudry, J. S., & Klavas, A. (1989, March). Survey of research on learning styles. *Educational Leadership*, 50-58.

Edelsky, C. (1990). Whose agenda is this anyway? A response to McKenna, Robinson & Miller. *Educational Researcher, 19*, 7-11.

Education and Skills Committee. (2004). *Teaching children to read*. Minutes of evidence taken before Education and Skills Committee, House of Commons, United Kingdom Parliament. Retrieved from http://www.publications.parliment.uk/pa/cm200304/cmselect/cmeduski/uc1269-i/uc126902.htm

Eldredge, L. (1991). An experiment with a modified whole language approach in first grade classrooms. *Reading Research and Instruction, 30*(3), 21-38.

Ferrari, J. (2012). A decade of lost action on literacy. *The Australian*, 22 December 2012, p.1.

Feuer, J. F., Towne, L., & Shavelson, R. (2002). Scientific culture and educational research. *Educational Researcher 31*(8), 4-14.

Fields, B. A., & Kempe, A. (1992). Corrective feedback in Whole Language teaching: Implications for children with learning problems. *Australian Journal of Special Education, 16*(2), 22-31.

Flesch, R. (1955). *Why Johnny can't read*. New York: Harper & Row.

Foorman, B. R. (1995). Research on 'the Great Debate': Code-oriented versus Whole Language approaches to reading instruction. *School Psychology Review, 24*, 376- 392.

Francis, W. N. (1965). *The English language: An introduction*. New York: Norton.

Gable, R. A., & Warren, S. F. (1993). The enduring value of instructional research. In Robert Gable & Steven Warren (Eds.), *Advances in mental retardation and developmental disabilities: Strategies for teaching students with mild to severe mental retardation* (pp. 1-7). Philadelphia: Jessica Kingsley.

Gersten, R., & Dimino, J. (1993). Visions & revisions: A special education perspective on the Whole Language controversy. *Remedial & Special Education, 14*(4), 5- 13.

Gersten, R., Keating, T. & Becker, W. (1988). The continued impact of the Direct Instruction model: Longitudinal studies of Follow Through students. *Education & Treatment of Children, 11*, 318-327.

Gibson, E., & Levin, H. (1978). *The psychology of reading*. MA: MIT.

Gonzalez, J. E. J., & Valle, I. H. (2000). Word identification and reading disorders in the Spanish language. *Journal of Learning Disabilities, 33*, 44-55.

Goodman, K. S. (1986). *What's whole in whole language*. Richmond Hill, Ontario: Scholastic.

Goodman, K. S. (1989, March). Whole Language is whole: A response to Heymsfeld. *Educational Leadership*, 69- 72.

Groff, P. (1990). An analysis of the debate: Teaching reading without conveying phonics information. *Interchange, 21*(4), 1-14.

Hammill, D. D. (2004). What we know about correlates of reading. *Exceptional Children, 70*, 453-469.

Harrison, B. (2002, April). Do we have a literacy crisis? *Reading Reform Foundation, 48*. Retrieved from http://www.rrf.org.uk/do%20we%20have%20a%literacy%crisis.htm

Hempenstall, K. (1996). The gulf between educational research and policy: The example of direct instruction and Whole Language. *Behaviour Change, 13,* 33-46.

Hempenstall, K. (1997). The role of phonemic awareness in beginning reading: A review. *Behaviour Change, 14*, 201-214.

Hempenstall, K. (2002). Phonological processing and phonics: Towards an understanding of their relationship to each other and to reading development. *Australian Journal of Learning Disabilities, 7*(1), 4-29.

Henry, M. K. (1993). The role of decoding in reading research and instruction. *Reading and Writing: An Interdisciplinary Journal, 5*, 105-112.

Heymsfeld, C. R. (1989, March). Filling the hole in Whole Language. *Educational Leadership*, 65-68.

Hirsch, E. D. (2003, Spring). Reading comprehension requires knowledge of words and the world. *American Educator*, 10-13, 16-22, 28-29, 48.

House, E. R., Glass, G. V., McLean, L. D., & Walker, D. F. (1978). No simple answer: Critique of the Follow Through evaluation. *Harvard Educational Review, 48*, 128-160.

Huang, H. S., & Hanley, R. (1994). Phonological awareness and visual skills in learning to read Chinese and English. *Cognition, 54*, 73-98.

Iverson, S., & Tunmer, W. E. (1993). Phonological processing skills and the Reading Recovery program. *Journal of Educational Psychology, 85*, 112-126.

Johnson, B., & Stone, E. (1991). Is whole language restructuring our classroom? *Contemporary Education, 62*, 102-104.

Kame'enui, E. J., & Simmons, D. C. (1990). *Designing instructional strategies: The prevention of academic learning problems*. Columbus, OH: Merrill.

Kavale, K. A. (1990). Variances & verities in learning disability interventions. In T. Scruggs, & B. Wong (Eds.), *Intervention research in learning disabilities* (pp 3- 33). New York: Springer Verlag.

Kavale, K. A., & Forness, S. R. (1985). The science of learning disabilities. San Diego: College Hill Press.

Kavale, K. A., & Forness, S. R. (1987). Substance over style: Assessing the efficacy of modality testing and teaching. *Exceptional Children, 54*, 228-239.

Kirsch, I., & Guthrie, J. (1984). Adult reading practices for work & leisure. *Adult Education Quarterly, 34*, 213- 232.

Klesius, J.P., Griffith, P.L., & Zielonka, P. (1991). A Whole Language and traditional instruction comparison: Overall effectiveness and development of the alphabetic principle. *Reading Research & Instruction, 30*(2), 47-61.

Liberman, I. Y., & Liberman, A. M. (1990). Whole Language vs. code emphasis: Underlying assumptions and their implications for reading instruction. *Annals of Dyslexia, 40*, 51-75.

Louden, W., Chan, L. K. S., Elkins, J., Greaves, D., House, H., Milton, M., Nichols, S., Rivalland, J., Rohl, M., & van Kraayenoord, C. (2000). *Mapping the territory – Primary students with learning difficulties: Literacy and numeracy* (Vols. 1-3). Canberra: Department of Education Training and Youth Affairs.

Marks, G. N., & Ainley, J. (1997). *Reading comprehension and numeracy among junior secondary school students in Australia*. Melbourne: Australian Council for Educational Research.

Mather, N. (1992). Whole Language reading instruction for students with learning disabilities: Caught in the cross fire. *Learning Disabilities Research & Practice, 7*, 87-95.

McCaslin, M. M. (1989). Whole Language: Theory, instruction, and future implementation. *The Elementary School Journal, 90*, 223-229.

McGee, L. M., & Lomax, R. G. (1990). On combining apples & oranges: A response to Stahl & Miller. *Review of Educational Research, 60*, 133-140.

McGinitie, W. H. (1991, March) Reading instruction: Plus ca change. *Educational Leadership*, 55-58.

McKenna, M. C., Robinson, R. D. & Miller, J. W. (1990a). Whole Language and the need for open enquiry: A rejoinder to Edelsky. *Educational Researcher, 19*, 12-13.

McKenna, M. C., Robinson, R. D., & Miller, J. W. (1990b). Whole Language: A research agenda for the nineties. *Educational Researcher, 19*, 3-6.

McKenna, M. C., Stahl, S. A., & Reinking, D. (1994). A critical commentary on research, politics, and Whole Language. *Journal of Reading Behavior, 26*, 211-233.

Meyer, L. A., Gersten, R. M. & Gutkin, J. (1983). Direct Instruction: A Project Follow Through success story in an inner-city school. *The Elementary School Journal, 84*, 241-252.

Moats, L. C. (2000). *Whole Language lives on: The illusion of balanced literacy.* Washington, DC: Thomas B.Fordham Foundation. Retrieved from *www.ldonline.org/article/6394/*

National Commission on Excellence in Education (1983). *A nation at risk: The imperative for educational reform*. Washington, DC: US Government Printing Office.

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, DC: U.S. Department of Health and Human Services. Retrieved from http://www.nationalreadingpanal.org

Newman, J. M. (1991). Whole language: A changed universe. *Contemporary Education, 62*, 70-75.

Paulesu, E., Demonet, J-F., Fazio, F., McCrory, E., Chanoine, V., Brunswick, N., Cappa, S. F., Cossu, G., Habib, M., Frith, C. D., & Frith, U. (2001). Dyslexia: Cultural diversity and biological unity. *Science, 291*, 2165-2167.

Pollack, M., & Pickarz, J. (1963). *Reading programs & problem readers*. New York: David McKay.

Primary National Strategy (2006). *Primary framework for literacy and mathematics*. UK: Department of Education and Skills. Retrieved from http://www.standards.dfes.gov.uk/primaryframeworks/

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. Hillsdale, NJ: Prentice Hall.

Reid Lyon, G. (2005). *Children of the Code Interview*. Retrieved from http://www.childrenofthecode.org/interviews/lyon.htm#WholeLanguage

Rosenshine, B. V., & Stevens, R. (1984). Classroom instruction in reading. In D. Pearson (Ed.), *Handbook of Reading Research* (3rd ed., pp. 376-391). New York: MacMillan.

Roth, S. F., & Beck, I. L. (1987). Theoretical & instructional implications of the assessment of two microcomputer word recognition programs. *Reading Research Quarterly*, 22, 197-218.

Schweinhart, L. J., Weikart, D. P., & Larner, W. B. (1986). Consequences of three pre-school curriculum models through age 15. *Early Childhood Research Quarterly, 1*(1), 15-45.

Shannon, P. (1994). *People who live in glass houses*. In C. B. Smith (Ed.), Whole Language: The debate (pp. 97- 120). Bloomington, IN: ERIC.

Share, D. L. (1995). Phonological recoding and selfteaching: Sine qua non of reading acquisition. *Cognition, 55*, 151-218.

Singer, H., & Balow, I. H. (1981). Overcoming educational disadvantageness. In J. T. Guthrie (Ed.), *Comprehension and teaching: Research review* (pp. 101-114). Newark, DE: International Reading Association.

Smith, F. (1978). *Understanding reading*. New York: Holt, Rinehart & Winston.

Smith, F. (1992). Learning to read: The never-ending debate. *Phi Delta Kappan, 74*, 432-441.

Smith, P. G. (1991). A practical guide to whole language in the intermediate classroom. *Contemporary Education, 62*, 88-95.

Smith, S. B., Simmons D. C., & Kame'enui, E. J. (1995). *Synthesis of research on phonological awareness: Principles and implications for reading acquisition*. NCITE Technical Report No 21. Retrieved from http://darkwinguoregon.edu/~ncite/reading/PhonSyn.html

Snider, V. E. (1992). Learning styles and learning to read: A critique. *Remedial and Special Education, 13*(1), 6-18.

Spear-Swerling, L., & Sternberg, R. J. (1994). The road not taken: An integrative theoretical model of reading disability. *Journal of Learning Disabilities, 27*, 91-103.

Spiegel, D. L. (1992). Blending Whole Language and systematic direct instruction. *The Reading Teacher, 46*(1), 38-44.

Stahl, S. A. (1990). Riding the pendulum: A rejoinder to Schickedanz, and McGee and Lomax. *Review of Educational Research, 60*, 141-151.

Stahl, S. A., & Kuhn, M. R. (1995). Does Whole Language or instruction matched to learning styles help children learn to read? *School Psychology Review, 24*, 393-404.

Stahl, S. A., & Miller, P. D. (1989). Whole Language & language experience approaches for beginning reading: A quantitative research synthesis. *Review of Educational Research, 59*, 87-116.

Stahl, S. A., McKenna, M. C., & Pagnucco, J. R. (1994). The effect of whole language instruction: An update and a reappraisal. *Educational Psychologist, 29*, 175- 185.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360-406.

Stanovich, K. E. (1993). The language code: Issues in word recognition. In S. R. Yussen & M. C. Smith (Eds.), *Reading across the life span* (pp. 111-135). New York: Springer-Verlag

Stanovich, K. E. (1994). Romance and reality. *The Reading Teacher, 47*, 280-291.

Stedman, L., & Kaestle, C. (1987). Literacy & reading performance in the United States, from 1880 to the present. *Reading Research Quarterly, 23*(1), 8-46.

Stone, J. E. (April 23, 1996). Developmentalism: An obscure but pervasive restriction on educational improvement. *Education Policy Analysis Archives*. Retrieved from http://seamonkey.ed.asu.edu/epaa

Strauss, S. L., & Altwerger, B. (2007). The logographic nature of English alphabetics and the fallacy of direct intensive phonics instruction. *Journal of Early Childhood Literacy, 7*(3), 299-317.

Sykes, S. (1991). A whole language perspective on reading & writing. *Australian Journal of Remedial Education, 23*(2), 23-27.

Topsfield, J. (2013). *Gillard ups the ante on schools*. The Age, Wed Feb 27).

Tunmer, W. E. & Hoover, W. A. (1993). Phonological recoding skill and beginning reading. *Reading & Writing: An Interdisciplinary Journal, 5*, 161-179.

Vellutino, F. R. (1991). Introduction to three studies on reading acquisition: Convergent findings on theoretical foundations of code-oriented versus whole language approaches to reading instruction. *Journal of Educational Psychology, 83*, 437-443.

Weaver, C. (1988). *Reading process & practice: From sociopsycholinguistics to whole language*. Portsmouth, New York: Heinemann Educational Books.

Weir, R. (1990). Philosophy, cultural beliefs and literacy. *Interchange, 21*(4), 24-33.

Wendorf, J. (2004). *Children of the Code Interview*. Executive Director, National Center for Learning Disabilities. Retrieved from http://www.childrenofthecode.org/interviews/wendorf.htm#NeurobiologicalDeficienciesVsInstruc

Yates, G. C. R. (1988). Classroom research into effective teaching. *Australian Journal of Remedial Education, 20*(1), 4-9.