

# 1 Introduction and overview

***Marian Sainsbury***

The nature of ‘reading’ is something usually taken for granted. In contemporary societies, the use of literacy for a vast range of social and personal purposes is so widespread that it is rarely questioned. Within the education system, reading becomes an explicit focus of attention, with substantial resources devoted to the teaching and learning of literacy. Even here, however, the definition of ‘reading’ is usually not discussed, although one can be inferred from the kinds of teaching and learning activities adopted. It is in the research community that the nature of reading becomes a defined area of study, and here, as will become apparent, there are major disagreements between different academic traditions over what is included and implied by the term ‘reading’.

Essentially, this book sets out to explore some of the theories, practices and conflicts that surround the idea of reading at the beginning of the 21st century. In order to do this, it adopts a particular perspective: that of assessment. Researchers, educationalists and the population at large have questions about how well people read. Often, though not exclusively, these people are children who are still in the process of mastering reading. This need to assess leads inevitably to the question ‘What exactly are the skills and understandings that we want to know about, in order to gauge reading ability?’ Thus it is that a particular definition of reading becomes made concrete in an assessment. By scrutinising tests and other instruments, it is possible to study the definition of reading – the construct – specified or assumed in each one. It is the existence of this concrete evidence in the form of tests and other instruments that makes assessment a promising springboard for investigating the nature of reading.

In 2003–4, a series of seminars was held in England, supported by the research funding body the Economic and Social Research Council, with the purpose of exploring the construct of reading. The participants were selected with the deliberate intention of allowing interaction between different disciplines, and consisted of a group of specialists in assessment and reading from the United Kingdom, France and the United States. There were cognitive psychologists with research interests in reading; educationalists with a range of research backgrounds in the teaching and learning of literacy and literature; and assessment specialists. Unusually in such gatherings, there was a strong representation of test developers, whose day-to-day research activities included the practical processes of devising, trialling and refining actual reading tests.

This group set out to bring together their varying experiences of and perspectives on the construct of reading. The seminars were open-ended and built in generous time for discussion, in recognition of the complexity of the subject matter. Each individual chapter in this volume explicates its reasoning and rationale, with references that situate it within its own research background. However, some ‘fault lines’ in the arguments can

be set out in general terms, and these apply both to ideas about reading and to ideas about assessment.

When we read, we consciously or unconsciously recognise written symbols as words with meaning. The act of reading includes deciphering, or decoding, written words and letters, transforming them into recognisable language, and understanding their meaning. Meaning is intricately tied up with communication, and communication of many kinds of meanings occupies a central role in human social intercourse. There is a fundamental divide between researchers who focus primarily on the decoding of words and those who focus primarily upon reading as an act of meaning-communication. For the former group, 'reading' proper is recognising the words; the uses of those words to communicate meaning and support social interaction are interesting, but not essential to the construct of reading. For the latter group, by contrast, it is not possible to make sense of the notion of 'reading' without communicating meanings; the communicative act is primary, and the specific skills involved in decoding written words cannot logically be separated from this.

These two perspectives can be thought of as competing paradigms: theory-systems that shape experience. The paradigm determines what counts as evidence, what observations are relevant, and even what is observed. Because the difference between them is logical, definitional rather than empirical, no observation can prove that one is right rather than the other. But this rather bleak view of paradigm competition does not rule out an understanding of both, nor a rapprochement between them in practice. In the real world of a literate society and an education system preparing children to participate in it, the stark differences outlined above are masked. All agree that children need to acquire the ability to recognise words fluently and to use this ability to facilitate and enrich their everyday lives.

At the same time, there are equally fundamental disagreements about what – and whom – assessment is for, and once again these can be seen as competing paradigms. One kind of assessment purpose is to pinpoint strengths and weaknesses in reading development and to diagnose barriers to that development. Such assessments give rise to indications that guide teaching or prescribe remedial action. These formative and diagnostic assessments can be seen as broadly for the benefit of the learner, but also of teachers and other professionals, whose effectiveness is enhanced by this information.

Formative assessment stresses the value of informative feedback in the course of ongoing teaching. The information obtained from informal, day-to-day assessment is used by the teacher to provide better-focused teaching. It can also be used by the learner as a powerful tool for improvement. If pupils are able to monitor their own learning in this way, rather than relying on the teacher and other outsiders, they can play an active part in planning their own learning experiences. In this classroom use, the assessment is very informal. The evidence can be entirely ephemeral, such as a pupil's answer to a teacher's question, or take the form of feedback comments on a pupil's written work. It is also possible to use more formal assessments in this formative way. Rather than focus on the numerical score obtained in a test, it is possible to make an analysis of

the strengths and weaknesses demonstrated at individual, group or class level, and to use this information to plan the curriculum.

Diagnostic assessment is used when a child is experiencing difficulties in learning to read, in order to pinpoint the perceptual or cognitive problems that underlie the lack of progress. For this purpose, carefully designed batteries of subtests are devised, and are administered on an individual basis by an educational psychologist.

An entirely different purpose for assessment is certification. Final examinations and tests assess the reading curriculum covered in the course of schooling. The certificates awarded on the basis of these assessments serve to attest to the competence and understanding of the student. They thus benefit the student, in providing a recognised measure of attainment, but also society, where they fulfil the purpose of selecting candidates for further study or for employment. They are high-stakes assessments, because the individual student's life chances are affected by them.

In some countries in recent years, however, notably the UK and the USA, the predominant purpose for assessment is political accountability. Governments have a legitimate interest in improving educational standards. Better national attainment in literacy benefits individuals as well as enhancing the economic performance of a country. In this context, tests have the role of providing the performance outcomes that are used by government and the public to evaluate progress towards defined targets. As a result, the tests acquire high stakes for the local authorities, schools and teachers who are being held to account for their pupils' performance.

The participants in the seminars represented a range of interests in and allegiances to these differing views on the nature of reading and on the purpose of assessment. They were brought together with the aim of understanding more about one another's perspectives, and perhaps finding an overarching position that brought them closer. In the latter aspiration, it is fair to say that the seminars had only limited success, as the incompatibility of the underpinning theories became if anything more evident. This will be discussed in more detail in the concluding comments of the book. But in the aim of fostering mutual understanding and respect, the seminar series can be counted a success.

This book is the outcome of those seminars, and the themes outlined above are worked out in a variety of cross-cutting ways in the following chapters. Part 1 is devoted to explicating in more depth some of the theoretical underpinnings of reading and of assessment. Each of the authors in this section sets out a single perspective; it is only later in the book that the links between them become apparent.

Marian Sainsbury starts this process by outlining the evolving theory of construct validity in assessment and suggesting an overall shape for the construct of reading that attempts to integrate competing points of view.

In John Beech's chapter, a psychological perspective is advanced that can broadly be situated in the tradition focusing primarily on the decoding of the written word. In highlighting the contribution of psychology to the study of reading, he stresses the value of soundly researched psychological theories, arguing that these are rarely if ever embodied in reading tests. His stance on assessment is to highlight its diagnostic

and theory-building potential, with tests contributing to research and to remediation of difficulties.

Alastair Pollitt and Lynda Taylor also adopt the perspective of cognitive psychology, and agree with Beech's argument that reading test design fails to reflect psychological research and theory. Their particular focus is upon the processes that characterise reading comprehension and they use research findings and practical illustration to build a picture of the process of constructing meaning from text. In this way, their chapter directly contributes to the delineation of the construct of reading that runs through the book.

In the fourth chapter of theoretical exploration, Colin Harrison takes a radically different perspective, situated firmly within socio-cultural and literary traditions. His starting point is the challenge posed to the definition of reading by postmodernism. On this view, meaning itself is shifting and ephemeral and texts are susceptible to a variety of readings, none of which is privileged over others. He draws from this position a demanding set of requirements for any valid assessment of reading. Harrison's chapter adds a new dimension to the emerging outline of the construct of reading and raises fundamental questions about assessment itself.

Ros Fisher adopts yet another stance, this time looking at reading and its assessment from the perspective of teachers and children. Her chapter argues strongly for the centrality of these stakeholders in the assessment of reading. For Fisher, purposes for assessment must be evaluated according to how well they support the enterprise of teaching and learning, and the construct of reading in tests is contrasted unfavourably with school literacy, which itself is narrower than the broader, richer and more subtle nature of literacy as social practice.

These introductory chapters highlight the assumptions and demands of several of the most influential paradigms that compete in the definition of reading and its assessment. By their very nature, these chapters present incommensurable theories; the arguments of one cannot be translated into the terms of another. However, any serious attempt to investigate the construct of reading must take account of this bewildering complexity, reflecting as it does the variety of viewpoints attracted by this important human activity. There is no attempt here to build bridges. Instead, these opening chapters raise and illuminate the issues that arise in different forms in the remainder of the book.

Any test or assessment can be seen as a concrete embodiment of a construct of reading and of an assessment purpose. In designing a test, in espousing an assessment system at any level from individual to class to nation, decisions are made about reading and about purpose which define a specific form in the context of a specific use. This book is about actual assessments just as much as the theories that underlie them. By examining reading tests and other less formal assessment approaches, the abstract notions set out in the first part of the book can be investigated in practice. The later chapters take on this task, sometimes focusing closely on a single test, sometimes more than one, sometimes following up practical implications of particular ideas.

In the second section, three chapters cast a historical light upon the assessment of reading, revealing patterns of change over time and discovering reasons and causes in

the historical context. P David Pearson surveys some key moments in the history of reading tests in the USA, pointing up the evolution in underlying theory that gave rise to each new development. Similarly, Chris Whetton highlights significant points in reading testing in the UK, but argues that political and social influences are at least as important as theoretical developments in determining key outcomes. Paul Thompson's chapter chronicles the course of one influential innovation in the UK, demonstrating in its own way the jostling of literary and educational theories with political imperatives in determining the shape of reading assessment at one moment in time.

Leading on from these historical insights, the third section of the book looks at the cutting edge of current work and finds two apparently contradictory ideas occupying equally prominent positions. On the one hand, developments in information and communication technology have led to a burgeoning interest in computer-based assessment. Colin Harrison introduces some fundamental issues and principles for consideration. Once again juxtaposing theory with practice, Claudia Leacock's chapter describes an innovative computer program, already operational, that makes possible the assessment of open written responses, releasing computer-based reading tests from the limitations of the multiple-choice question.

On the other hand, a contrasting area of interest and innovation that can be discerned in current thought is the use of informal assessment by both teacher and pupils to support learning – known broadly as the 'assessment for learning' movement. Gordon Stobart sets out the principles of formative classroom assessment and applies them to reading. To complement this, Lorna Pepper, Rifat Siddiqui and Andrew Watts describe a research project investigating the value of giving feedback in a specific form to students who have taken a reading test.

In implementing national assessment systems, governments make decisions about the nature of reading and the purpose of assessment that set the agenda for national discourse. The seminar group included participants who were directly involved in devising the national assessments in England, Wales, Scotland, France and the USA. The insights from these very different systems make up the fourth and final part of the book.

Marian Sainsbury and Andrew Watts describe a system of national testing in England that attempts to combine a complex, meaning-centred, literary construct of reading with the constraints of a high-stakes testing regime. Roger Palmer and David Watcyn Jones describe a similar construct of reading operating in Wales, but their chapter traces the evolution of a similar high-stakes accountability assessment system into one that supports teachers in assessing for formative and diagnostic purposes. This perspective is further amplified by Louise Hayward and Ernie Spencer, writing about Scotland. Here, there is an established commitment to formative assessment that is worked out in all aspects of the national assessment system.

The national evaluations in France take a distinctive view of both the construct of reading and of the nature and purpose of national assessment. Martine Rémond describes a set of formal national tests that are entirely formative in purpose, and that embody a definition of reading which accords more importance to grammatical knowledge than is usual in the Anglo-Saxon world. Finally, Patricia Donahue sets out yet

another different national response to the set of questions posed by the definition and purpose of reading assessment. The National Assessment of Educational Progress in the USA is a national survey of reading attainment that yields indications of performance that are crucial to political decision-making but low stakes for individual students. The construct of reading can broadly be situated within a 'responsive reading' paradigm. These five national assessment systems therefore represent a variety of ways in which contemporary societies obtain their evidence about reading, and demonstrate how these governmental decisions are both reflections and determinants of national values.

Each of these chapters is capable of standing alone, giving a summary and overview of a particular perspective. The book can be used for reference, bringing together a collection of theoretical and practical information about the assessment of reading in its political, educational, geographical, historical and contemporary contexts. Reading the entire book brings out the interaction between these factors, as principles are juxtaposed with concrete examples, political demands with academic, social influences with individual, theories with classrooms.