Assessment Framework for the 2004 National Assessment of English Reading (NAER):

First and Fifth Classes

Assessment Framework for the 2004 National Assessment of English Reading (NAER):
First and Fifth Classes

Judith Cosgrove
Lisa Milis
Gerry Shiel
Patrick Forde
Hannah Wardle

Educational Research Centre

Further copies of this report may be downloaded from http://www.erc.ie/na2004.html
For more information on the National Assessment of English Reading, email na2004@erc.ie

Table of Contents

PREFACE

This document describes the framework and specifications for the sixth National Assessment of English Reading (NAER 2004). The content of NAER 2004 reflects a balance between the Primary school English curriculum, previous national assessments, and recent developments in international assessments of reading. Thus, sources used to guide the development of this framework include the current Primary School English Curriculum, and recent international assessments of reading literacy (notably the Progress in International Reading Literacy Survey, or PIRLS, and the Programme for International Student Assessment, or PISA). This document represents a collaborative effort between researchers at the Educational Research Centre, and representatives on the NAER National Committee. It was felt important to publish a framework for NAER since for the first time, pupils at both first and fifth class levels are to be assessed. All previous national surveys of reading in Ireland assessed pupils in fourth or fifth classes only. A second reason for publishing the framework is to encourage and advance discussion about the many facets of reading literacy, especially in the light of the current Primary School English Curriculum, which was introduced into schools in 1999. A survey of mathematics at fourth class level will be carried out at the same time as NAER 2004. However, this framework document describes the assessment of reading only.

We would like to acknowledge the a number of individuals who contributed to the development of this report. Thanks to Professor Al Beaton, Boston College, for his expert advice on scaling, test design, and the monitoring of trends in reading achievement. Thanks to Margaret Dowling (Primary teacher, test developer) for writing some of the test material for the first class tests, and to Eithne Kennedy (St Patrick's College, Dublin) for advice on the appropriateness of the new fifth class test material. Thanks to David Millar (ERC) for developing the main survey sample design. Thanks to Deirdre Hackett (ERC) for her work on scaling the pilot survey data. Thanks to Mary Rohan (ERC) for continued administrative support. Thanks to the NAER national committee for guiding the direction of this framework, the content and focus of the background questionnaires, and aspects of the survey administration. Thanks to the principals, teachers and pupils in the three schools who assisted us in pre-piloting the test materials. This process enabled us to make important refinements to the content and layout of the tests before piloting them. Finally, many thanks to the Inspectors of the Department of Education and Science for overseeing the administration of the pilot survey which took place in May 2003, and the principals, class teachers, Learning-Support teachers, pupils, and their parents/guardians in the 32 schools in Dublin, Cork and Limerick who took part in the pilot survey. Without their participation and commitment, NAER 2004 could not happen.

Timely and widespread dissemination of results are important if a survey of academic achievement is to be both effective and relevant. Therefore, it is planned to publish the initial outcomes of NAER 2004 in 2005. Two publications are planned:

- A short, highly accessible summary report, including an executive summary, which will detail the main outcomes and policy implications of the survey. This summary will be widely disseminated and made available on the Educational Research Centre's website.
- A longer, more detailed analysis of the outcomes, including some additional technical information about the survey's design and analysis. The publication of this more scholarly report may be accompanied by a series of workshops explaining outcomes in detail, and allowing time for discussion and feedback.

BACKGROUND TO THE 2004 NATIONAL ASSESSMENT OF ENGLISH READING

The 2004 National Assessment of English Reading (NAER 2004) is the sixth of a series of national assessments, dating from 1972. All previous reading assessments examined the achievements of pupils in fourth/fifth classes (Cosgrove, Kellaghan, Forde, & Morgan, 2000). The 2004 survey is a departure from these in one major respect: the reading skills of pupils at first class level are to be assessed for the first time, in addition to assessing pupils at fifth class level.

A further development since 1998 is the inclusion of new assessment material at fifth class level. In both 1993 and 1998, pupils in fifth class were assessed using the same reading assessment, called Tasks for the Assessment of Reading Achievement (TARA 93/98), developed specifically for large scale national assessments of reading (Martin, Forde & Hickey, 1991). TARA 93/98 consists of five test booklets which give broader domain coverage than would a single assessment booklet administered to all pupils. However, some passages are now somewhat dated and two of the five booklets are to be retired, and replaced with more up-to-date material that reflects the content of the Primary English Curriculum (NCCA, 1999), and recent international assessments of reading literacy, especially the Progress in International Reading Literacy Survey (PIRLS; Campbell et al., 2001) and the Programme for International Student Assessment (PISA; OECD, 1999).

Achievements of pupils in the 1993 and 1998 assessments were scaled using Item Response Theory (IRT), with an overall mean in 1993 set at 250 and the standard deviation at 50. Achievements of pupils were reported as a total scale score, which was a weighted mean of three text type subscales: narrative, expository (continuous informational texts) and documents (non-continuous informational texts).

Cognitive measures like TARA 93/98 can indicate the reading standards of the population as a whole. However, this information is limited in that it can only tell us about the overall reading standards without reference to the contexts in and conditions under which pupils' achievements occur. Hence NAER 1993 and 1998 also included a number of contextual measures derived from questionnaires given to principals, teachers, parents and pupils, in order that variations in achievement might be linked to pupils' home, school and social contexts, as well as individual differences in their dispositions and habits. The 2004 assessment follows this structure in that both cognitive and contextual measures are to be administered, and the remainder of this document describes these measures.

The work on NAER is guided by a national committee, membership of which may be found in Appendix A.

The aims of NAER 2004 are:

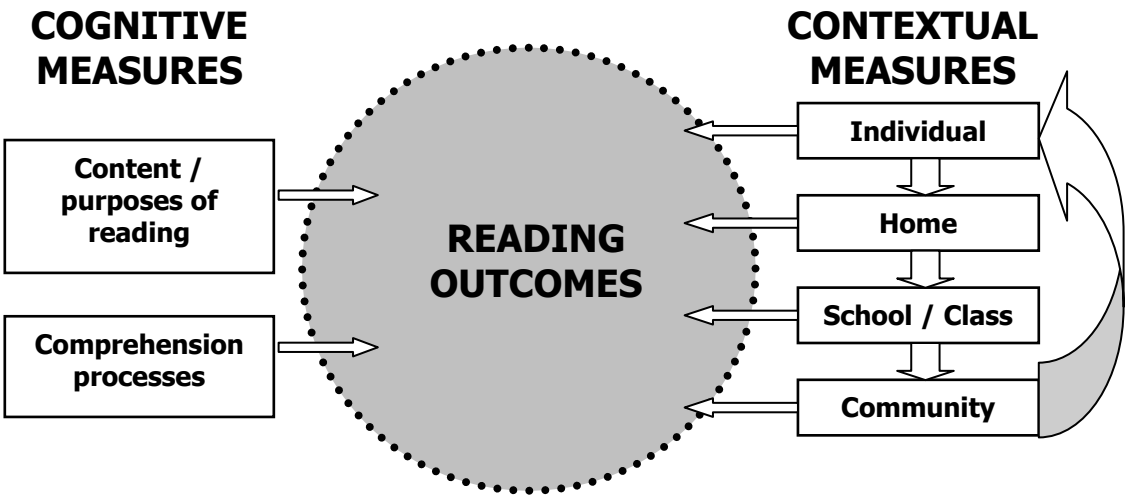
- To establish current reading standards of first and fifth class pupils;
- To compare outcomes at fifth class level with the outcomes of NAER 1998;
- To provide high quality and reliable data for the Department of Education and Science to assist in policy review, formulation and resource allocation related to English reading;
- To examine school, teacher, home background, and pupil factors which may be related to reading standards;
- To provide a basis with which to compare future assessments of English reading.

OVERVIEW OF ASSESSMENT FRAMEWORK

The specification of the NAER 2004 cognitive measure of reading distinguishes between comprehension processes, in which readers engage, and the content of a text, i.e., the purpose for which it is written. The contexts in which reading takes place, and how these relate to variations in achievement, are also described in this framework. The contexts are described in terms of their proximity to the child's immediate environment (Figure 1).

The remainder of this document describes reading processes and reading contents/purposes, and the measures used to assess them in detail. Where relevant, comparisons are made with the Primary School English Curriculum, the 1993 and 1998 national assessments, and assessment frameworks underlying recent international assessments of reading.

Figure 1 Framework for NAER 2004: Contextual and cognitive measures



TEST SPECIFICATION

Definition of Reading

Wixson and Peter's (1984) definition of reading was adopted for NAER 1993 and 1998:

> *Reading is the process of constructing meaning through the dynamic interaction among the reader's existing knowledge, the information suggested by the written language, and the context of the reading situation.*

This definition reflects contemporary views of reading in the sense that it goes beyond simple decoding and enabling skills and conceptualises reading as a dynamic, interactive, and constructive process. However, more recent conceptualisations of reading have also emphasised functions of and purposes for reading (see Figure 2).

Figure 2 Definitions of reading used in international assessments

IEA Reading Literacy Survey (RLS) (1991) (ages 9 and 14):

> *…the ability to understand and use those written language forms that are required by society and/or valued by the individual* (Elley, 1994, p. 5).

International Adult Literacy Survey (IALS) (ages 16-65):

> *…using printed and written information to function in society, to achieve one's goals, and to develop one's knowledge and potential* (Murray, Kirsch & Jenkins, 1998, p. 17).

Programme for International Student Assessment (PISA) (age 15):

> *…understanding, using, and reflecting on written texts, in order to achieve one's goals, to develop one's knowledge and potential, and to participate in society* (OECD, 1999, p.20).

Progress in International Reading Literacy Study (PIRLS) (age 9):

> *…the ability to understand and use those written language forms required by society and/or valued by the individual. Young readers can construct meaning from a variety of texts. They read to learn, to participate in communities of readers, and for enjoyment* (Campbell, Kelly, Mullis, Martin & Sainsbury, 2001, p.3).

References to *using* texts as well as understanding them are present in all the definitions in Figure 2. Although the IEA Reading Literacy Survey (IEA RLS) and International Adult Literacy Survey (IALS) definitions focus on the functional aspects of literacy[1], they are not as comprehensive as the definitions adopted for the Programme for International Student Assessment (PISA) and Progress in Reading Literacy Survey (PIRLS). The PISA and PIRLS definitions, in addition to emphasising uses of and attitudes towards reading, also make reference to the processes and skills of comprehension. Because PISA focuses on the experience of 15-year olds, who are typically preparing to enter the workforce or higher education, the definition emphasises 'participation in society'. PIRLS, on the other hand, is more concerned

[1] The term 'functional aspects of literacy' is used in the framework document to refer to the usefulness and purposes of reading, and should not be confused with 'functional texts' such as newspapers and catalogues.

with the reading experiences of 9-year olds and the acquisition of reading skills. Therefore, the PIRLS definition emphasises reading to learn, reading for recreation, and 'participation in communities of readers' (e.g., home and classroom) rather than in 'society'. Since Wixson and Peter's definition did not make explicit reference to the functional aspect of reading, it is expanded for NAER 2004, using part of the PIRLS definition as follows:

> *Reading is the process of constructing meaning through the dynamic interaction among the reader's existing knowledge, the information suggested by the written language, and the context of the reading situation. Young readers read to learn, to participate in communities of readers, and for enjoyment.*

This definition is intended to convey the notion that readers interact with the text in the context of a particular reading experience to actively construct meaning. It also emphasises the importance of using reading for a variety of purposes. In the PIRLS framework, two main purposes are identified: reading for literary experience, and reading for information. In reading for literary experience, "the reader brings to the text his or her own experiences and perspectives, and engages with the text to become involved in imagined events, settings, actions, consequences, characters, atmospheres, feelings and ideas, and to enjoy language itself" (Campbell et al., 2001, p. 17). In reading to acquire and use information, by contrast, the reader "engages with aspects of the real world, rather than with imagined worlds. Through information texts, readers can acquire knowledge about the world and why things work as they do, and can use this knowledge in reasoning and in action" (Campbell et al., 2001, p. 17).

The definition of reading proposed for NAER 2004 is also consistent with the Primary School English Curriculum, which asserts that "the ultimate objective of reading is comprehension or the reconstruction of meaning. The meaning… grows gradually and in the process is redefined, revised and reformulated by the reader when he/she engages in reading the text and in reflecting on it" (NCCA, 1999, p. 61). The curriculum also emphasises the functional nature of reading, noting that the acquisition of reading skills "is central to the effective learning in every area of the curriculum and to the child's social and community life outside school" (NCCA, 1999, p. 26).

Reports on national assessments of reading conducted by the National Foundation of Educational Research (NFER) for England and Wales do not include a formal definition of reading; however, a description of the content of the reading tests is provided and these suggest a content breakdown similar to that of NAER 2004. In the assessment of 9-year olds (Brooks, Pugh, & Schagen, 1996), the Reading Ability Series Level B (Kispal, Gorman, & Whetton, 1989) was used, which contains a narrative and an expository test, totalling 35 items split evenly between the two domains, with 12 open-ended and the remainder multiple choice in format. In an assessment of reading of 8-year olds in England and Wales (Brooks, Schagen, & Nastat, 1998), the Reading Ability Series Level A was used. The test consists of 25 items, also a mixture of multiple-choice and open-ended questions. Three passages, two non-continuous (a notice and a menu) and one continuous literary, comprise the test material. No information on reading processes is included in these reports.

Content and Processes of Reading

Specifying the reading task

Introduction

In NAER 2004, a content-by-process classification that is similar to that used in TARA 93/98 is used, whereby the content categories of the reading task (i.e., types of passages to be read) are cross-classified with the types of process involved in

reading. However, the way in which these two aspects of reading are conceptualised has been adapted to reflect more recent views on reading, as evidenced in recent international assessments and the Primary School English Curriculum.

Content in TARA 93/98
To ensure that the materials selected for the assessment are representative of the reading activities of first and fifth class pupils and appropriate to their interests and abilities, it is important to specify the reading content for pupils at these class levels. Before a description of the TARA 04 reading content is given, a description of the content of TARA 93/98 is provided, since TARA 04 builds on this work.

In the course of developing TARA 93/98, six classifications of text types were arrived at by examining the form (i.e., temporal and/or topical) and function (i.e., to affect readers' emotions, behaviours, or knowledge) of text. Six functions were identified: narration, lyric, directions, persuasion, temporal exposition, and topical exposition. Based on a review of six basic senior primary school textbooks carried out in the course of developing TARA 93/98, two major categories of text, narrative and expository, and two minor categories of text, poetry and reference (documents), were identified and formed the basis of the subdomains of reading within TARA 93/98. However, while documents items contributed to descriptions of pupils' reading performance, poetry items did not, since (i) there were not sufficient items to provide a reliable description of pupil performance in this domain, and (ii) while poetry is included in the 1971 and 1999 Primary School English Curricula, it was felt that poetry pertains more to oral than to reading skills.

Content in TARA 04
The documents subdomain of TARA 93/98 comprised primarily reference texts (dictionary excerpts, tables of contents, timetables). The info-non-continuous (documents) subdomain in TARA 04 is broadened for both first and fifth classes with reference to the distinction made between continuous and non-continuous texts which was used by Kirsch and his colleagues (e.g., Kirsch, 2001) in the course of the developing assessment used in the International Adult Literacy Survey, and also evident in PISA (e.g., Kirsch et al., 2002), which examines the reading skills of 15-year olds. Kirsch distinguishes between continuous and non-continuous texts as follows:

> *Continuous texts are typically composed of sentences organised into paragraphs. These may fit into larger structures such as sections, chapters, and books. The organisation of non-continuous texts is different from continuous texts and so allows the reader to employ different strategies for entering and extracting information. Most frequently, non-continuous texts are organised in matrix format, based on combinations of lists.* (2001, p.13)

This suggests that aside from reference materials and timetables, non-continuous texts also include functional information texts such as forms, advertisements, diagrams, and maps. Therefore the documents component of the assessment in 2004 includes such texts, in addition to reference materials.

Table 1 outlines the different text types which fall into the purpose/structure cross-classification for NAER 2004 for first and fifth classes. The resulting content categories are in most respects similar to those used in TARA 93/98. Narrative texts fall within the continuous texts/reading for literary experience category, while expository texts fit into the category of continuous texts/reading for informational purposes. Non-continuous texts which are written for informational purposes include documents (reference materials).

This classification is also consistent with the Primary School English Curriculum, which classifies text as either expository, where the principal function is to inform and

explain; narrative, where the text is mainly concerned with telling a story; or diagrammatic/representational (documents), where the text is designed to present and illustrate information.

The Primary School English Curriculum thus recognises the need for pupils to engage with a range of texts as they progress through school. With respect to the continuous text/reading for literary experience classification, the curriculum indicates that the telling and reading of stories should be a feature of children's experience at every class level. The curriculum suggests that in first/second class, the child should be able to explore texts such as stories and picture books, while fifth/sixth class pupils should read a range of narrative text, including stories, myths, legends, novels, and plays.

The Primary School English curriculum also emphasises that children should be given the opportunity to engage with a varied range of informational texts. For example, in fifth/sixth class, this includes learning about the structure and function of parts of a newspaper, such as editorials and news features. The curriculum for first/second class states that the child should engage with a wide variety of texts including informational material.

Table 1 Classification of texts in TARA 04

Text structure	**Purpose for which text was primarily written**	
	Reading for literary experience*	Reading to acquire and use information*
Continuous	Narrative Poetry**	Narrative Descriptive Exposition Argumentation / persuasion
Non-continuous	-	Reference materials Forms Advertisements Charts and graphs Diagrams/Schematics Tables and matrices Maps Lists Injunctive / instructive

Note. The text types listed in each cell are based on the classification adopted in the PISA framework (OECD, 1999).
*Fiction is generally associated with reading for literary experience while non-fiction is generally associated with reading to acquire and use information, as indicated in the Primary School English Curriculum.
**Although included in the TARA 93/98 test, poetry did not contribute to pupils' scaled achievement scores.

The Primary School English curriculum indicates that children should also be given an opportunity to explore a variety of representational and diagrammatic text. At first/second class level, this includes invitations, cards, charts, and lists. Children at these class levels should also be able to perform simple information retrieval tasks using a table of contents and a simple index. At fifth/sixth class, pupils should explore various parts of a newspaper, including advertisements and TV and radio programme schedules. They should also learn to use reference materials, be able to read and interpret texts such as forms, menus, and timetables, and be able to find information using texts graphs, flowcharts, and texts with pictorial and diagrammatic data.

The Primary School Curriculum also includes reading in a range of subjects (History, Geography, Science, Drama, Mathematics, Social and Personal Health Education).

Therefore, texts in NAER 2004 at first and fifth class levels cover a broad range of topic areas.

Processes in TARA 93/98
In addition to specifying the form and functions of text, the assessment framework addresses the processes of comprehension, as suggested in the definition of reading literacy. Because TARA 04 builds on the specifications of TARA 93/98, a description of these processes is given first.

In the TARA development project, three process categories were described, and a description of reading outcomes with respect to these processes was included in the report on NAER 1993 and NAER 1998 (Cosgrove et al., 2000). The processes specified were:

- *Local*—where the reader is required to construct meaning from one, two, or three contiguous sentences. Literal questions are typically employed, although simple inferences might sometimes be required.
- *Text-wide*—where the reader is required to integrate information from across the text, other than that specified in Local. Such questions require inferential comprehension.
- *Text-plus*—the reader must rely on information from his/her own experience as well as in the text. Inferential comprehension and/or evaluation is required.

These classifications broadly parallel Bloom et al.'s (1956) more traditional taxonomies of levels of cognitive functioning: literal, inferential and evaluative.

In recent years, researchers and test developers have attempted to specify the underlying processes in greater detail and recent frameworks have described ways of viewing, understanding or responding to texts. For example, the IEA RLS assessed six skills: verbatim match, paraphrase information, locate main theme, make an inference, locate information in a document, and follow directions. In IALS, four levels of processing were identified: locating, cycling, integrating and generating. However, neither of these studies reported performance by process area. PISA identified five processes, though for reporting purposes these were collapsed into three: retrieve information (forming a broad understanding and retrieving information); interpret information (developing an interpretation); and reflect on/evaluate information (reflecting on the content and form of texts). The PIRLS framework describes four types of comprehension processes: retrieve information, make straightforward inferences, interpret and integrate ideas and information, and examine and evaluate aspects of the text. PIRLS, however, has not, to date, reported achievement in terms of these reading processes; rather, reading achievement was reported in terms of purpose (literary experience and information acquisition) (Mullis et al., 2003).

Processes in TARA 04
Although the process classifications in TARA and PIRLS are not identical, there is a high degree of similarity between the four types of processes identified in PIRLS and the skills categories described in TARA 93/98 (literal/inferential/interpretative/evaluative processes and local/text-wide/text-plus processes). Thus, it was decided to elaborate on the process categories for TARA 04, using the process classification in PIRLS, so as to reflect the increased emphasis given to processes underlying reading comprehension. Since PIRLS focused on the acquisition of reading among 9-year olds, it was felt that this process classification was more developmentally appropriate than PISA for both first and fifth classes (with some modifications at each class level). A description of the four types of comprehension processes adopted from PIRLS (Campbell et al., 2001, pp. 10-14), and the relationship of each to the processes identified in the NAER 93/98 framework (Cosgrove et al., 2000), follows.

It should be noted that the process of integrating information has two meanings: (i) the combining of information in the text with opinion or prior knowledge, and (ii) the combining of different pieces of information residing in the text. Since these two integrative processes often co-occur in making inferences or interpreting text, we use the term here in both senses.

The first type of process, *focusing on and retrieving explicitly stated information,* requires little or no inferring or interpreting as the required answer is explicitly stated in the text. However, it does require that the reader understand what is stated in the text, as well as how that information relates to the information sought. The focus typically remains at the sentence or phrase level. This process corresponds closely to the literal/local skills categories associated with TARA 93/98.

The second process, *making straightforward inferences,* is also text-based in that the answer, though it may not be explicitly stated, resides in the text (i.e., can be inferred directly from the text). The information necessary for making the inference is located within the text, but the relationship between two or more pieces of information may not be explicitly stated and must therefore be inferred. The reader may be required to focus on local meaning, or on more global meanings representing the whole text. This process may be compared to either the literal/local categories in TARA 93/98 (if the reader is required to focus on a meaning residing in part of the text) or the inferential/text-wide categories specified in TARA 93/98 (if the reader is required to focus on meanings across the text as a whole).

The third process, *interpreting and integrating ideas and information*, requires the reader to process and integrate text. This process may require a focus on either local or global meanings, or may require readers to discern main ideas and overall themes in the text. In this process, readers often integrate personal knowledge and experience with meaning that resides in the text in order to construct a more complete understanding of the text. Integrating ideas corresponds, to some extent, to the inferential/text-wide process, and interpreting ideas and information can be linked to the evaluative/text-plus categories of TARA 93/98, since it requires the reader to draw from his/her own experience and knowledge of the world.

In the fourth type of processing, e*xamining and evaluating content, language, and textual elements*, the focus shifts from constructing meaning to standing apart from the text and evaluating it, either from a personal perspective, or from a more critical and objective viewpoint. This type of comprehension process corresponds to the evaluative/text-plus categories described in TARA 93/98, since the reader must draw from their own experience and knowledge of the world to evaluate the text.

Figure 3 shows some examples of the four types of text processing specified in TARA 04. For the sake of brevity, the four processes are referred to as retrieve, infer, interpret, and evaluate in this document.

Figure 3 Examples of the four reading processes in TARA 04

Examples of Retrieval Processing:

- Looking for specific ideas, definitions or phrases.
- Identifying the setting of a story (e.g., time, place).
- Finding the main theme of the text (when explicitly stated).

Examples of Straightforward Inferential Processing:

- Inferring that one event caused another event.
- Concluding what is the main point made by a series of arguments.
- Identifying generalisations made in the text.
- Describing the relationship between two characters.

Examples of Higher-Level Interpretive and Integrative Processing:

- Discerning the overall message or theme of a text.
- Considering an alternative to actions of characters.
- Comparing and contrasting text information.
- Inferring a story's mood or tone.
- Interpreting a real-world application of text information.

Examples of Evaluative Processing:

- Evaluating the plausibility of the events described.
- Describing how the author devised a surprise ending.
- Judging the completeness or clarity of information in the text.
- Determining an author's perspective on the central topic.

Processes in the Primary School English Curriculum

The classification of items into these four broad processes is also consistent with the Primary School English Curriculum. In first/second class, the emphasis is placed on the first two processes. For example, the curriculum states that first/second class pupils should be able to recall details and events, assimilate facts, and retell stories. These processes are associated with the retrieval of explicitly stated information and making straightforward inferences.

According to the curriculum, first/second class pupils should also be able to respond to characters and events in a story, explore different attitudes and feelings by imagining what it would be like to be certain characters, and be able to give a considered personal opinion of a book in oral or written form. These tasks correspond to the processes of interpreting and integrating information, and, to some degree, evaluating content. However, since the majority of the focus at first/second class is on retrieving and integrative processes, the first class test for NAER 2004 (TARA 04 1st) does not contain any items that require evaluative processing.

In fifth/sixth class, the curriculum tends to focus on the two higher processes, since the majority of pupils at this age have learnt to read and are now reading to learn. The curriculum suggests that pupils at this level should be able to:

- use comprehension skills such as analysing, confirming, synthesising and correlating to aid deduction, problem-solving, and prediction;
- support arguments and opinions with evidence from the text;
- relate personal experience to ideas and emotions conveyed in the text;
- distinguish between fact and opinion, and bias and objectivity, in the text; and
- examine similarities and differences in various types of text.

The first task reflects a mixture of the second, third and fourth processes and a general problem-solving approach to reading. The second and third tasks correspond to the interpreting and integrating information process category, while the final two of these tasks entail the process of evaluating the content and structure of the text.

Thus a comparison of the relative emphases that the Primary School English Curriculum accords first/second classes and fifth/sixth classes indicates that a test at first class level should focus more on retrieval and inference processes, while at fifth class level, the emphasis should be on interpretative and evaluative processes, also ensuring adequate coverage of retrieval and inference processes to describe the reading skills of more basic level readers.

It should be noted that the curriculum also refers to a range of *enabling skills* that pupils should develop in each of the class levels. For example, at first/second class level, pupils should:

- build a sight vocabulary of common words;
- increase their awareness of sounds;
- learn to connect the beginning of words and syllables with their rhyming parts;
- learn about common word endings, word families and roots of words; and
- use their knowledge of letter-sound relationships, grammar, syntax and surrounding text in attempting to identify unfamiliar words.

Fifth/sixth class pupils should, for example, achieve proficiency in word identifications and develop their ability to recognise and understand words using root words, prefixes, suffixes and syllabication.

Although these enabling skills are critically important, the curriculum recognises that the ultimate objective of reading is *comprehension*. Thus, the focus of TARA 04 is on assessing comprehension processes rather than enabling skills. However, the Teacher Questionnaires do include questions that cover instruction in enabling skills.

Item formats

TARA 93/98 comprised 249 items, all of which were traditional multiple choice in format. However, recent international assessments have emphasised the importance of tapping reading skills through a variety of formats. PISA, for example, distinguishes between five items types (OECD, 2001):

- Traditional multiple choice, where the respondent chooses one of (usually) four possible options.
- Complex multiple choice, where the respondent selects one of each of a series of (usually) binary response options (e.g., true/false).
- Short response, where the answer consists of a single word, number, or short phrase.
- Closed constructed response, where there is some degree of breadth in the range of possible answers, but the answer itself is quite succinct.
- Open constructed response, where the answer may comprise several phrases or sentences, and where there is a broad range of possible correct responses.

Open response item formats are particularly appropriate for the interactive constructive nature of reading as indicated in the definition of reading adopted for NAER 2004. They are also well-suited to assessing higher-level interpretative, evaluative skills, particularly where pupils need to draw on their own experiences, and/or where a wide range of interpretation and responses is possible (Campbell et al., 2001). Therefore, about one-third of the new items at fifth class level (TARA 04 5th) are open response items. Consistent with the PIRLS framework, the only distinction made is that between open response and multiple choice items (the finer distinction made in PISA relates primarily to marking requirements and the degree of judgement required in assigning marks). At first class level (TARA 04 1st), it was felt

that the level of development of some pupils' writing may not accurately reflect their level of reading development and therefore potentially confound a measure of reading ability, so all items were designed to be multiple choice in format. Multiple choice item formats have the advantage of lower rates of missingness, are not confounded with pupils' writing speed, and guessing can be controlled for in the scaling process.

Mapping The Domain Specification Onto The Assessment

Once the broad content and process specifications have been described, the next issue to consider is the amount of weight within the test that should be given to each content type (purpose) and each reading process. Figure 4 presents the interaction of the two dimensions: each of the four processes of comprehension will be assessed within each of the three different contents.

The use of a two-way classification, where the content categories of the reading task (e.g., types of passages to be read) are cross-classified with the types of process involved in reading, is consistent with the Primary School English Curriculum, which suggests that a range of comprehension processes are to be used with each text type (NCCA, 1999, p. 62). For reasons of brevity, reading for literary experience/continuous is referred to literary-continuous, reading to acquire and use information/continuous is referred to as info-continuous, and reading to acquire and use information/non-continuous is referred to as info-non-continuous throughout the remainder of this document.

Figure 4 TARA 04—Cross-classification of reading content/purpose, and processes of comprehension

	Content (purpose of reading by text structure)		
Processes of comprehension	*Reading for literary experience /continuous text*	*Reading to acquire and use information /continuous text*	*Reading to acquire and use information /non-continuous text*
Focus on and retrieve explicitly stated information			
Make straightforward inferences			
Interpret and integrate ideas and information			
Examine and evaluate content, language, and textual elements			

Because the English Curriculum does not specify the relative amounts of reading in these areas, an analysis of English textbooks for first and fifth class was undertaken. The analysis is described in full in Appendix B and leads to the content breakdown shown in Table 2. The percentages are given as ranges rather than discrete numbers since the framework is intended to act as a guideline rather than an absolute

specification. It can be noted, however, that at fifth class level, the percentage of items devoted to the three content types is comparable to TARA 93/98, in which 40% of items were narrative, 40% expository and 20% documents. The corresponding figures for 2004 are 40%, 30%, and 30%. This is also comparable to PIRLS, in which 50% of items were devoted to reading for literary purpose, and 50% to reading to acquire and use information. At first class level, just over half of the items are associated with narrative texts, about 25% with expository, and 20% with documents.

Table 2 TARA 04—Desired percentages of reading assessment devoted to content and processes

Content (Purpose)	**First class**	**Fifth class**
Reading for literary experience / continuous	50-60	35-45
Reading to acquire and use information / continuous	20-30	25-35
Reading to acquire and use information / non-continuous	15-25	25-35
Processes of comprehension		
Focus on and retrieve information	40-50	25-35
Make straightforward inferences	30-40	25-35
Interpret and integrate ideas and information	15-25	15-25
Examine and evaluate content, language, and textual elements	0	15-25

It is evident that the proportion of items in each content and process category differs between first and fifth class. This is because of the different age-groups targeted. This is consistent with the Primary School English Curriculum, which emphasises basic processes more in the lower class levels. Further, since narrative/story texts comprise a large part of the reading experiences of first class pupils, the first class assessment places more emphasis on reading for literary experience than the fifth class assessment.

In TARA 93/98, approximately 48% of the items were devoted to local processes, 36% to text-wide processes, and 16% to text-plus processes (Table 3). Retrieving information corresponds to the local classification, while making straightforward inferences corresponds to both local and text-wide processes. Thus, devoting 60% or so of items in NAER 2004 to these two processes at fifth class level is broadly keeping in line with TARA, while also entailing a slight reduction of emphasis on this type of processing in favour of interpretative and evaluative processes. Interpreting and integrating ideas corresponds to the text-wide and text-plus classifications, and evaluating corresponds to the text-plus classification. Thus, devoting about 40% of items to these two processes is a slight increase compared to TARA 93/98. This slight shift in emphasis is in line with both the revisions to the national curriculum and recent international assessments. For example, in PIRLS, the highest proportion of items were in inferring and interpreting categories (30% each for making straightforward inferences and interpreting/integrating information), while the remainder of items were evenly divided between retrieval type and evaluate type processes (20% each). At first class level, there is a greater emphasis on retrieving information (45%) and less on infer (35%) and interpret (20%) processes.

Overall, the proposed distribution of items across processes for TARA 04 5th finds the middle ground between TARA 93/98 and PIRLS (e.g., slightly more emphasis is placed on the higher processes in TARA 04 5th than was the case in TARA 93/98 and slightly more emphasis is placed on the two lower processes in TARA 04 1st than in PIRLS). The proportion of items tapping each process in PISA 2000 are include in the table. Since first class pupils are still learning to read, there is more emphasis on retrieval and inference processes in than on the more complex interpretative and evaluative processes in comparison with PIRLS. At fifth class level, while approximately 20% of items tap evaluative processes, there is a somewhat higher emphasis on retrieval and inferential processes compared with PISA.

Table 3 Percentages of items devoted to comprehension processes in TARA93/98, PIRLS and PISA, and desired percentages for TARA 04

Process	**TARA 04 1st (desired)**	**TARA 04 1st (desired)**	**TARA 93/98 (actual)**	**PIRLS (desired)**	**PISA (actual)**
Retrieve	40-50	25-35	48 (local)	20	30 (retrieve)
Inference	30-40	25-35	36 (textwide)	30	50 (interpret)
Integrate/ interpret	15-25	15-25		30	
Evaluate	0	15-25	16 (textplus)	20	20 (reflect/evaluate)

Outcomes of the Pilot Survey

Overview

Prior to the main assessment in May 2004, the draft tests, questionnaires and administrative procedures were trialled in a pilot survey of first and fifth class pupils in May 2003. While all material at first class is new and required piloting (and comprised six booklets of material), three of the five TARA 93/98 test booklets are retained in NAER 2004 at fifth class level. Therefore, only the new material was piloted at fifth class level (in three test booklets). The purposes of the pilot survey were to gauge the appropriateness of the tests at both first and fifth classes in terms of difficulty and length; eliminate problematic items due to gender bias, extreme difficulty/easiness, or other psychometric problems; gauge the appropriateness of the questionnaire measures; conduct exploratory analyses of the questionnaire and test outcomes; and evaluate and refine administrative procedures. The main outcomes with respect to the tests are detailed below. The content of the questionnaires is discussed in a later section.

Background

The NAER 2003 pilot took place during the last two weeks in May. Administration of tests and questionnaires was carried out by classroom teachers, and overseen by Inspectors of the Department of Education and Science. Inspectors were briefed on the aims and procedures of the assessment in three regional briefings about one week prior to the assessment, and provided with all assessment materials. Teachers were provided with detailed information booklet on the aspects of the survey administration, including a 'script' for administering the tests and questionnaires. Both Inspectors and teachers were encouraged to give feedback on the assessment; teachers were given a test session rating form to record their comments on the timing and content of the tests and questionnaires. This feedback, together with the data itself, was used to guide revisions to the tests and questionnaires.

Sample

The sample design was a convenience sample of 32 schools located in Dublin, Cork and Limerick.

Within the three counties, a number of schools were excluded for logistical/practical reasons:

- Schools participating in the Survey of Reading Literacy in Designated Disadvantaged Schools (SRL DDS), and replacement schools for that survey, which took place just before the NAER pilot (N=95, with 79 replacement schools)
- Schools that were included in a planned extension of the SRL DDS (N=39)
- Schools that participated in the pilot study of the SRL DDS earlier in 2003 (N=10)
- Senior schools that participated in the Breaking the Cycle evaluation in May 2003 (N=22)
- Schools with an enrolment size of less than 100 (N=1692)
- Schools not located in Dublin, Cork and Limerick.

Thus after applying these criteria to the original sampling frame of 3109 schools, 442 schools were eligible for selection.

Schools were divided into three explicit strata based on size as follows:

- Small schools: total enrolment 100-150 (N=109).
- Medium schools: total enrolment 151-250 (N=149).
- Large schools: total enrolment 251 or more (N=184).

Within strata, schools were further categorised using a number of implicit stratifying variables as follows:

- Location: Dublin (N=235), Cork (N=144) or. Limerick (N=63).
- Language: English (N=400), All Irish (N=37) or Gaeltacht (N=2).
- Designated disadvantaged status: Yes (N=43) or No (N=399).
- Percent female pupils in the school: None (N=73), 1 to 50% (N=190), 50% to 99% (N=123), 100% (N=56).

At first class level 1200 pupils were required to generate reliable item statistics, since six booklets were to be piloted, and each item requires about 200 responses. At fifth class level, three booklets were piloting, necessitating a minimum sample size of 600. The estimated required number of schools for first class, taking absences into account and assuming an average of 40 pupils or two classes per school, is 30. However, when the initial sample of 30 pupils was drawn, it was discovered that the expected total number of pupils was too low to supply reliable item statistics after adjusting for non-response. Therefore, two additional large mixed sex schools in the Dublin area with pupils at first class but not fifth class were drawn.

Schools for the first class sample were drawn first since more pupils at first class were required, taking care to select schools in proportion to the number of eligible schools in each of the three counties, and ensuring a balance in terms of language of instruction, designated disadvantaged status and gender composition of the school. A subset of 20 of the sampled schools was selected for the survey at fifth class level, taking care to maintain a similar balance in school characteristics in the sample. Table 4 details the characteristics of the sample at each class level compared with the population.

Table 4 Sample design for NAER pilot survey (May 2003)

	First Class	Fifth Class	Population
		Percent	
*Size**			
Small	12.5	10.0	33.1
Medium	40.6	45.0	36.6
Large	46.9	45.0	30.3
*Location***			
Dublin	53.1	50.0	45.1
Cork	31.3	30.0	39.2
Limerick	15.6	20.0	15.8
Designated			
No	81.3	80.0	90.0
Yes	18.8	20.0	10.0
Irish Language			
English	90.6	90.0	92.4
All Irish	6.3	5.0	4.1
Gaeltacht	3.1	5.0	3.5
% Female			
0%	12.5	10.0	8.6
1-50%	18.8	15.0	51.3
51-99%	56.3	60.0	35.2
100%	12.5	15.0	5.0
		Mean (SD)	
Average school size	304.6 (183.6)	302.0 (194.5)	135.0 (125.4)
Average N	40.3 (28.5)	39.4 (24.3)	**1st:** 16.3 (19.1) **5th:** 17.5 (18.9)
Average % female	54.0 (27.6)	53.4 (25.5)	47.3 (20.3)

*Population percentages for stratum/size do not include schools with an enrolment of less than 100 pupils.
**Population percentages for location includes all schools on the sampling frame in the three counties listed with an enrolment size of greater than 100 and does not include schools outside these areas.

It is desirable that the sample selected for the pilot survey broadly matches the population in terms of size, gender composition, language of instruction, and designated disadvantaged status. It should be noted that while the sample at both class levels is broadly representative of the population and has a balanced gender composition, it does under-represent small schools and, as a corollary, schools in sparsely populated/rural locations. This is to be addressed in the main survey sample design, whereby schools are not excluded on the basis of size; also, under-representation of certain school types can be redressed by applying sampling weights during analyses. It should also be noted that one Gaeltacht and two all Irish schools were included in the sample at first class level, and one Gaeltacht and one all Irish school at fifth class level, following concerns expressed by the NAER national committee that pupils at first class levels in these school types may be at a disadvantage on a test of English reading.

Within each school, all pupils at first class were invited to participate, and in 20 of the 32 selected schools, all pupils at fifth class were invited to participate. In seven of the schools, questionnaires were administered in addition to the tests, to pupils at both class levels, their parents, class teachers, Learning-Support teachers, and principals. While all pupils in ordinary classrooms were to be encouraged to take part in the assessment, some pupils were exempt. These included pupils with less than one year's instruction in English, pupils with moderate to severe learning disabilities, and pupils with a physical disability that would prevent them from participating.

Participation rates

All schools agreed to participate, yielding a school level response rate of 100%. All participating pupils attempted a NAER pilot test booklet and linking test, while pupils in seven schools completed questionnaires, as did their teachers, principals, and Learning-Support teachers. The linking test at fifth class was the 20-item General Reading Test (also used in NAER 93/98) and the linking test at first class was the 20-item Drumcondra Sentence Reading Test Form C (DSRT), consisting of items taken from DSRT Forms A and B (Educational Research Centre, 2002).

Table 5 shows the participation rates for the NAER pilot. Rates are quite high, close to 90% for all measures. In total, 1301 pupils were selected at first class. Of these, 16 were exempted and 12 returned completely blank booklets; the latter were counted as absent. Of the remainder, 1148 attempted the assessment. This yields a response rate of 89.3% and an exemption rate of 1.2%. Of participating students, 44.2% were male and 9.7% were attending designated disadvantaged schools. Note that the population percentage of 1st class pupils in designated schools in 2001-2002 was somewhat higher at 15%. The main reason for this difference is, as noted previously, that the survey of literacy in disadvantaged schools was concurrent, so many larger designated schools were not included in the list of schools eligible for the NAER pilot.

In total, 790 pupils were selected at fifth class. Of these, 13 were exempted. Of the remainder, 712 attempted the assessment. This yields a response rate of 91.6% and an exemption rate of 1.7%. Of participating students, 44.0% were male and 10.3% were attending designated disadvantaged schools. The population percentage of 5th class pupils in designated schools was similar to 1st class (15%).

Unlike the tests, the questionnaires were administered in just 7 of the schools. In these seven schools, at first class level, 222 were selected, 197 participated, 20 were absent and 5 were exempted. At fifth class level, 251 were selected, 225 participated, 20 were absent and 6 were exempted.

Table 5 Participation rates for each measure at first class and fifth class: NAER pilot survey

First Class	TARA 04 Pilot Test	Link Test	Pupil Questionnaire	Parent Questionnaire	Pupil Rating Form
N	1148	1148	197	191	199
%	89.3	89.3	90.8	88.0	91.7
Fifth Class					
N	712	712	225	218	225
%	91.6	91.6	91.8	89.0	91.8

Note. Response rates are calculated using the number selected minus the number exempted as the denominator.

Analysis of test data

General approach

Several sources were used to guide item analysis and item selection. These include technical documentation associated with the most recent reading (1998) and mathematics (1999) surveys (Cosgrove et al., 2000; Shiel & Kelly, 2001), as well as technical documentation from recent international studies such as PISA and TIMSS (e.g., Adams & Wu, 2002; Martin & Kelly, 1996). Other sources of information taken into account include the fit of retained items to the specification of the framework and test design, comments of Inspectors and teachers, and comments of the individuals that marked the open-ended test items at fifth class level.

The analysis had four broad phases.

1. The first comprised a review of all comments made by teachers and inspectors about the test.
2. The second comprised classical item analysis in ITEMAN 3.5 (Assessment Systems Corporation, 1995), where items were flagged as being potentially unsuitable based on a number of criteria that have been identified in the literature. These are:

- The relationship (point biserial correlation) between the correct response on a specific item and overall performance on the test is low (that is, the item doesn't 'fit' with what the test measures; the cut point was $\leq$.25)
- For multiple choice items, the relationship between an incorrect response (distractor) and overall performance on the test is positive (i.e., $\geq$.05, where the number of respondents is greater than 10 for that distractor)
- the item has an extreme percent correct (less than 10% or greater than 90%) – since such items tell us little about the population in general.

3. The third phase comprised of analysis and scaling of the data using item response theory (IRT) in BILOG (Mislevy & Bock, 1990) and BILOG-MG (Zimowsky, Muraki, Mislevy, & Bock, 1996), where there are three analysis stages. The second and third stages are of interest, where information about item characteristics and student ability is given. The following was examined:

- scaled item difficulty or threshold (which is similar but not quite the same as the percent correct since it takes pupil ability into account; extreme values $\geq$2 were flagged)
- item discrimination or slope (i.e. how well the item distinguishes between high and low ability readers; this was examined visually by looking at the item characteristic curves)
- guessing parameter for multiple choice items (if a distractor is very plausible, the guessing parameter will be higher)
- fit between the theoretical and actual ability distribution as indicated by item characteristic curves and, in the case of fifth class, a chi-square test.

It is common for the IRT analysis to reveal items that have already been flagged as problematic in the classical item analysis, so the IRT serves as an extra check in the

process. A review of IRT output is also important since it is planned to scale the tests using IRT in the full scale survey in 2004.

Several additional checks were also carried out. These comprised:

- Examination of differential item functioning by gender (i.e., taking overall ability into account, does a significant gender difference on individual items remain?).
- Examination of non-response (e.g., is there sufficient time for pupils to finish the test? Are there big differences in non-response between boys and girls, pupils in designated and non-designated schools, between multiple choice and constructed response items?)
- Computation of agreement between markers for open response items at fifth class level.
- Checks to see if partial credit marking was appropriate for some constructed response items at fifth class level.
- Checks for group equivalence (if test booklets were distributed at random, there should be no significant difference in the performance of pupils by TARA 04 pilot test booklet).
- Comparisons of the performance of subgroups (gender, designated disadvantaged status).
- Review of all items for correctness of process classification.
- Comparison of pupil ability and item difficulty estimates.

4. The retaining or dropping of an item was not an automatic process. In each case where an item was flagged it was carefully examined for logical and other flaws. Thus the fourth phase consisted of a series of meetings, using the information from the analyses listed above, to decide on a final item pool, to correct any logical flaws with retained items where possible, and to ensure that the selected items fit with the framework specifications.

Before presenting some characteristics of the final item pool, a brief summary of the main outcomes of the analyses is presented.

Summary of main outcomes

Differential item functioning: At first class, seven items, five favouring boys and two favouring girls, emerged as significant. These were examined and two dropped, one favouring boys, and one favouring girls. At fifth class, five items, two of them favouring boys, emerged. Three were dropped (two of them favouring girls). Differential item functioning analyses will also be conducted on the main study item pool before scaling the data.

Missingness—first class: Overall mean missingness, about 5%, is low, indicating that the timing and length of the test is a appropriate for first class pupils as a whole. Missingness in designated schools is somewhat higher (10% compared to 4%) although the difference is comparable to that for boys versus girls (7% compared to 3%). Missingness is higher for items appearing on the same page as the stimulus text and also towards the end of sections, but all values are below 10% with one exception. The booklets have been reformatted to ensure that the first question appears on the page facing the text. The differential missingness rates are not statistically significant except in the case of booklet B for gender.

Missingness—fifth class: Overall missingness is about 4% with little difference between boys and girls (4% for boys compared with 3% for girls). The difference is larger for designated and non-designated schools (11% compared with 3%). Missingness is higher for constructed response items than MCQ items (7% compared with 4%). The differential rates of missingness were statistically significant

for gender only in the case of booklet 1 for constructed response items, and for designated status for both item types in booklets 1 and 3.

Agreement between markers: At fifth class level, 80 of each booklet was marked twice using a rotated marking scheme similar to that specified in TIMSS 1995 (Mullis & Smith, 1996). The average percent agreement between marks assigned was 92% (range = 73-100%). The item with the lowest agreement (73%) has been dropped from the item pool, and feedback from markers has been used to clarify the marking schemes associated with some of the items. The agreement rate of 92% indicates a more than satisfactory level of agreement between markers (Mullis & Smith cite the figure of 85% for satisfactory marker agreement).

Partial credit: In the first instance, 18 of the constructed response items at fifth class level were marked using a partial credit, or coded into different categories, to distinguish any differences that might occur between pupils giving different type of response. These were then analysed in the scaling programme BILOG and the ability estimates of each item score or code group compared. In all cases it was deemed appropriate to collapse the categories into simple right-wrong scoring for scaling purposes; however, many of these categorical codes will be retained in the initial marking phase so as to provide additional useful qualitative information at the item level. It is not envisaged that partial credit scoring will be incorporated into the scaling process following the main study.

Group equivalence: It is essential, if a rotated booklet design is used, that each subsample, for each group of pupils taking each booklet, is equivalent. A comparison of the mean scores on the linking tests provide an indication of group equivalence. At both class levels, there was no difference in the average performance of the groups, thereby demonstrating group equivalence.

Representativeness of the sample: Since the first class performance on the test was higher than expected, there is a possibility that the sample is not representative of the population in general (especially since it is a convenience sample). Percent correct on the linking tests were therefore compared to percent correct for those of other known representative samples.

At first class, percent correct on the 20 DSRT items was compared with percent correct on the DSRT standardisation sample from 2002. The NAER pilot sample had a slightly higher mean of 64% compared with the DSRT mean (60%). We cannot test for significance between means since the NAER DSRT items are taken from two different forms. There are slightly more females than males in the NAER sample, however (56% of the sample was female) so this may account for some or all of the difference.

At fifth class, the most recent reference group available to compare performance on the link test is the 1998 NAER sample. Fifth class pupils in that survey had a mean percent correct of 68%. The NAER pilot sample had a mean percent correct of 71%, therefore the samples are quite similar on this ability measure.

However, given that the percentage of pupils in the NAER pilot sample in designated schools was lower than the population percentage, the final item pool has a mean percent correct that is a little higher than the ‘ideal’ percent correct of 65 to allow for the fact that the main survey sample may include a higher proportion of lower achieving pupils.

Combined standardised score: Subsequent to a number of meetings to select the final item pool, taking the aforementioned analyses, a desired average score[2] of about 65 and fit to the framework into account, booklets at each class level were placed on the same scale. This means that their ability scores as generated in BILOG (mean=0, sd=1.0) are comparable across booklets at each class level. These scores were then transformed to have a mean of 100 and standard deviation of 15. These scores were sent in October 2003 to teachers as feedback on pupil performance, accompanied by information on standard deviations and percentile ranks to assist interpretation.

Gender differences in overall performance: Using the standardised scale score (all booklets together) as the outcomes, girls in first class significantly outperform boys. The difference between the mean of girls (103.4) and boys (95.7) is about half a standard deviation. The magnitude of the overall difference for the TARA 04 pilot is about the same as that for the linking test (0.4 sd), indicating that the TARA 04 pilot test performs in the same way as the linking test in this regard. At fifth class level, the performance of girls is slightly higher than boys for the TARA 04 pilot and also for the link test but neither of these is significant. In fact the overall mean score of boys on TARA 04 (99.0) is just two points lower than that of girls (100.7); the mean difference for the linking test is less than 0.1 sd. This is an interesting finding, especially given that 15-year old female students in all countries that participated in PISA 2000 significantly outperformed males, and the gender difference in Ireland (29 scale points, or .31 sd; Shiel et al., 2001), which merits closer examination in NAER 2004.

Designated disadvantaged status and overall performance: Pupils in first class in non-designated schools significantly outperform those in schools designated as disadvantaged. The magnitude of the mean score difference for the TARA 04 first class pilot test data is about the same as that for the first class linking test. At fifth class level, the performance gap is also significant. The magnitude of the mean score difference for TARA 04 5th class pilot test data is also about the same as that for the fifth class linking test. Therefore, the new TARA tests at both class levels operate in the same way as the linking tests in this regard. Note, however, that these comparisons are based on pupils in just six designated disadvantaged schools and 26 non-designated schools.

Reclassification of items: In the initial classification of items, the distinctions between inference and interpret, and between interpret and evaluate, were not always clear. The classification guidelines were subsequently refined following a review of the PIRLS documentation, and all items were re-examined using the new guidelines, which allowed an item to be evaluative only if it entailed critically examining the content or structure of the text, i.e., distancing oneself from the text. This resulted in more interpret and fewer evaluate items at fifth class, and a reclassification of several items at first class, but not in any systematic direction. The TARA 93/98 items at fifth class level were also reclassified from the older process categories of local, textwide and textplus using the same guidelines as for TARA 04 5th.

Test targeting: Test targeting gives an indication of how appropriate the test difficulty level is for the population in question through a comparison of the distribution of the ability estimates with the item difficulty estimates. If the distributions are similar, then the test is well-matched to the population (e.g., Adams & Cartensen, 2002). A comparison of the standardised item difficulties and pupil ability estimates (scaled to have a mean of 0 and sd of 1) at first class level (Figure 5) indicate that there is a

[2] This is the 'ideal' percent correct for a test with all multiple choice, as mentioned, we have allowed it to go slightly higher to allow for the slight under-representation of pupils in designated schools. Since about one-third of items at fifth class are constructed response and the 'ideal' percent correct for these is lower at 50% (as guessing cannot occur), the percent correct for fifth class is a little lower at about 62%.

wide spread for both item and ability, and that the match between the two is quite good. However, the test may be slightly too easy for the target group. The reason for this is that the link test items as well as the TARA 04 1st items have been scaled together to produce the ability estimates due to the small number of items per pupil. This will not be the case in the main survey where TARA 04 1st items only will be used. A comparison of the difficulty and ability estimates at fifth class level (Figure 6) indicate that the test is well-suited to the population.

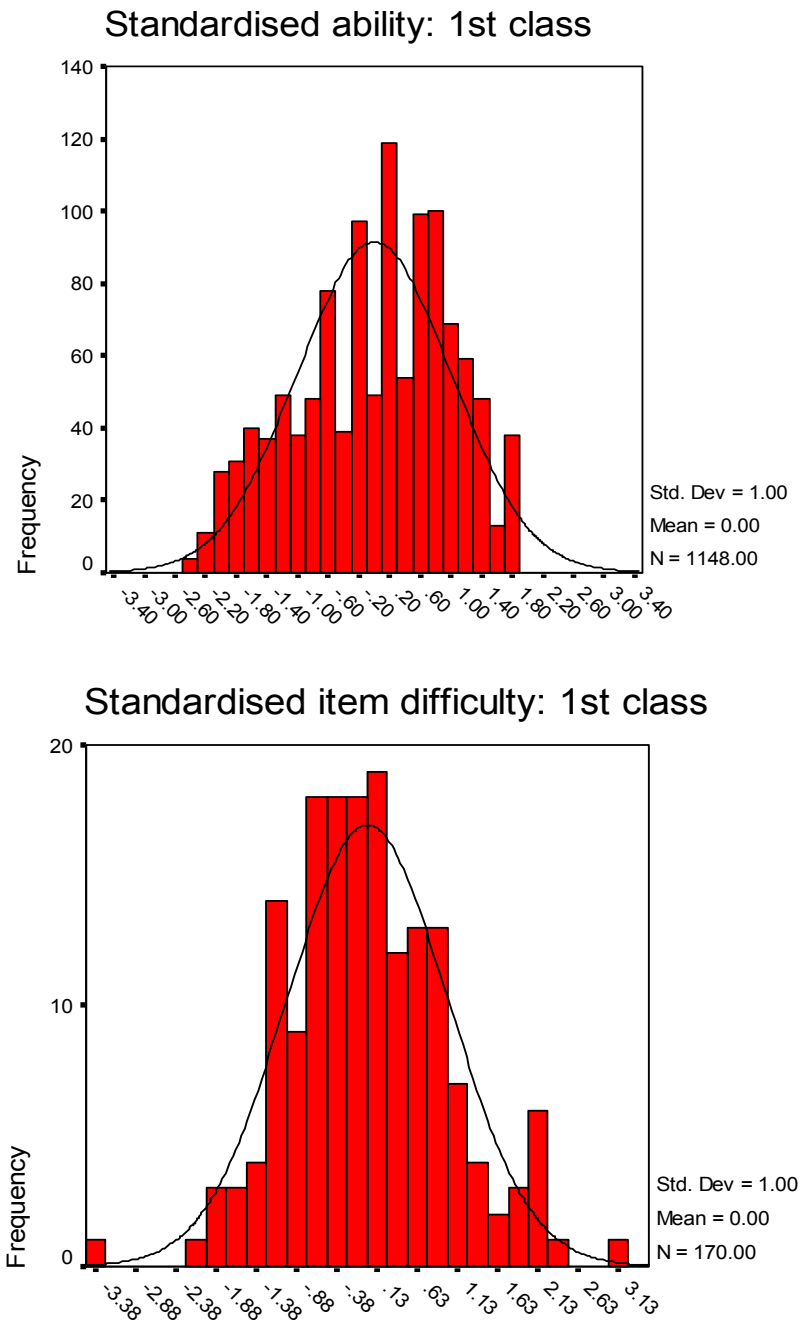


Figure 5 Comparison of item difficulty and pupil ability: First class

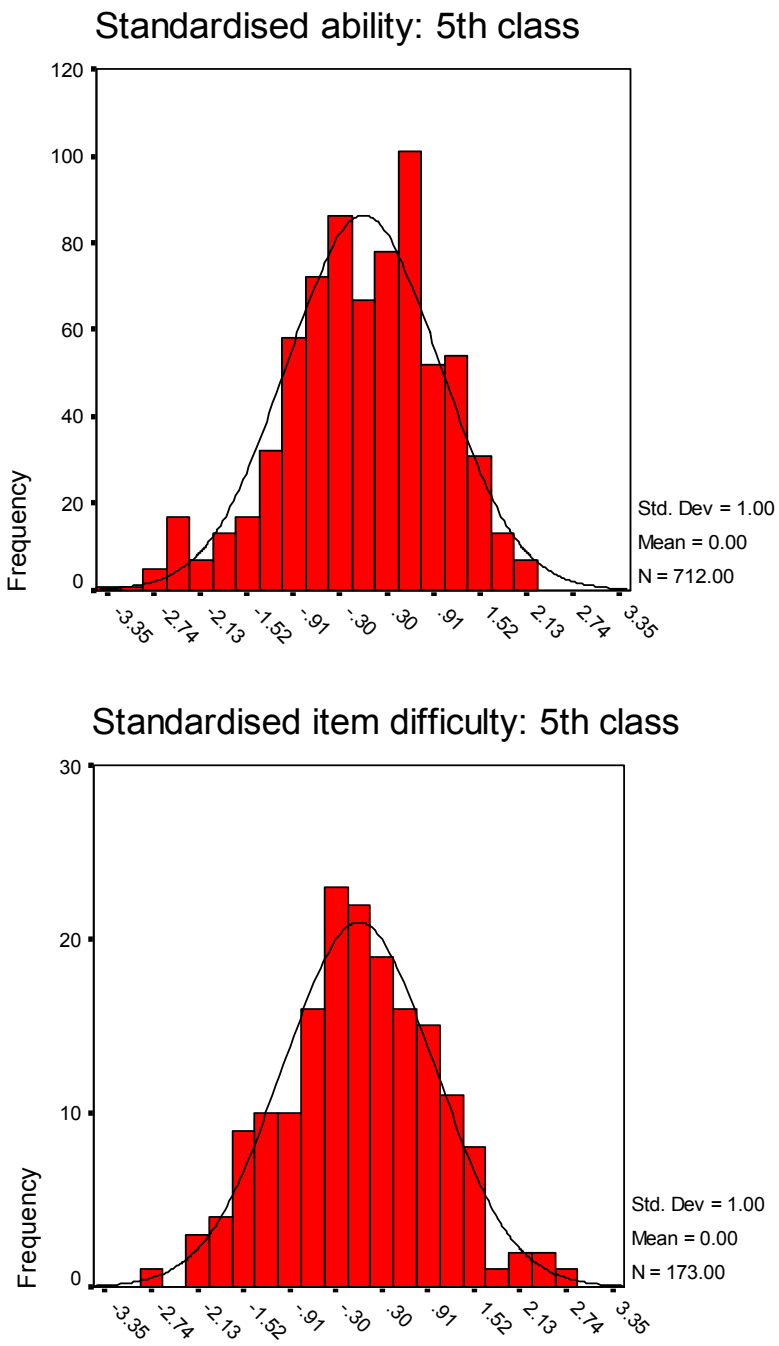


Figure 6 Comparison of item difficulty and pupil ability: Fifth class

Characteristics of Final Item Pool and Test Design

First class

The original item pool consisted of 169 items spread across six booklets. Table 6 shows the breakdown of these items by text type and reading process. Of these 169 items, 19 or 11% were found to have poor or unsatisfactory psychometric characteristics and were dropped from the item pool leaving 150 items intact. Appendix C contains first class some sample passages and items that were included in the pilot survey.

Table 6 Breakdown of items by process and text type, original item pool, first class

Process	N	Percent	Text Type	N	Percent
Retrieve	82	48.5	Literary	93	55.0
Infer	60	35.5	Info-Cont	35	20.7
Interpret	27	16.0	Info-Non-cont	41	24.3
Total	169	100.0	Total	169	100.0

The booklet design for the main survey for first class is shown in Table 7. It consists of nine 10- or 11-item blocks, and nine three-part booklets. Each block appears three times, once in each position of a booklet. Five of the blocks are literary, two are expository (informational continuous text), and two are documents (informational non-continuous text). Each block requires 10 minutes to complete and there is a short break after the first and second blocks.

The advantage of this design is that the estimates of item difficulty are not affected by the position of the items in the booklet. A second advantage is that the repetition of each block across three booklets results in a lower number of pupils in the required sample size (a minimum recommended sample size is 900-1,000 per item, which implies that a sample of about 3,000 pupils is sufficient). The blocks are ordered as Literary, Info-Continuous, Literary, Info-Non-continuous, Literary etc. to avoid having (i) all Literary items in a booklet, and (ii) more than two information blocks in a booklet. The blocks are relatively balanced in terms of reading load, number of items and item difficulty.

Table 7 Test design for TARA 04 1st

	Block/Position		
Booklet	***1***	***2***	***3***
1	1L	2IN	3L
2	2IN	3L	4IC
3	3L	4IC	5L
4	4IC	5L	6IN
5	5L	6IN	7L
6	6IN	7L	8IC
7	7L	8IC	9L
8	8IC	9L	1L
9	9L	1L	2IN

Note. L=literary, IN=information-non-continuous, IC=information-continuous in the Block IDs.

This design entailed the selection of a subset of the passages available, and the writing of a number of new items to satisfy the requirements of the framework. Thus 77 of the 150 available items were selected, and 18 new items were written. It was not deemed necessary to prepilot these new items since there are comparatively few

of them, and also since they are spread across nine passages, a large number of pupils would be required in the prepilot, making it inefficient. They will be evaluated carefully before scaling and some can be dropped if necessary. The breakdown of the final item pool of 95 items by text type and process, and mean percent correct of the 77 items for which these data are available is shown in Table 8. Retrieve items are easier than infer or interpret items, and informational-continuous items are somewhat easier than literary or informational-non-continuous items.

Table 8 Breakdown of items by process and text type, final item pool (first class; N=95; percent correct based on 77 items)

Process	N	Percent	Mean % correct	Text Type	N	Percent	Mean % correct
Retrieve	44	46.3	72.2	Literary	53	55.8	68.2
Infer	29	30.5	67.3	Info-Cont	20	21.1	73.5
Interpret	22	23.2	62.2	Info-Non-cont	22	23.2	66.7
Total	95	100.0	68.8	Total	95	100.0	68.8

Fifth class

The original item pool consisted of 222 items spread across three booklets. Table 9 shows the breakdown of these items by text type and reading process. Of these 222 items, 50 or 23% were found to have psychometric problems and were dropped from the item pool. Appendix C contains some fifth class sample passages and items that were included in the pilot survey.

Table 9 Breakdown of items by process and text type, original item pool (fifth class)

Process	N	Percent	Text Type	N	Percent	Item Type	N	Percent
Retrieve	80	36.0	Literary	73	32.9	MCQ	149	67.1
Infer	74	33.3	Info-Cont	79	35.6	Open	73	32.9
Interpret	58	26.1	Info-Non-cont	70	31.5			
Evaluate	10	4.5						
Total	222	100.0	Total	222	100.0	Total	222	100.0

The booklet design for the main survey for fifth class is shown in Table 10. Two of the five existing TARA 93/98 are to be dropped, and replaced with two new booklets of similar design. The design of the two new booklets is consistent with the TARA 93/98 design, consisting of eight non-overlapping blocks, two in each booklet half. An overlapping design is preferable, but not possible since the new booklets must match the design of the 93/98 booklets. The blocks are ordered as Literary, Info-Non-continuous, Info-Continuous, Info-Non-continuous in one booklet and Info-Continuous, Info-Non-continuous, Literary, Info-Non-continuous in the other. Booklets and booklet halves are relatively balanced in terms of reading load, number of items and item difficulty. Each booklet half takes 40 minutes to complete.

Table 10 Test design for TARA 04 5th, new booklets

	Block/Position			
Booklet	1	2	3	4
A2	IC1	IN1	L1	IN2
C2	L2	IN3	IC2	IN4

Note. L=literary, IN=information-non-continuous, IC=information-continuous in the Block IDs.

This design entailed the selection of a subset of the passages available, and the writing of a number of new items to increase the low number of evaluative items. Thus 116 of the 222 available items were selected, and 19 new items, almost all evaluative, were written. The breakdown of the final item pool of 135 items by text type and process, and mean percent correct of the 116 items for which these data are available along with a breakdown by process, text and item type, is shown in Table 11. Retrieve items are easier than infer, interpret, or evaluate items, and literary items are somewhat harder than informational-continuous or informational-non-continuous items. However some of the new literary items are expected to be easier than average, and it is hoped that this will redress the balance.

Table 11 Breakdown of items by process and text type, final item pool (new items only, N=135, percent correct based on 116 items)

Process	Percent	Mean % correct	Text Type	Percent	Mean % Correct	Item Type	Percent	Mean % Correct
Retrieve	31.1	70.3	Literary	35.6	57.7	MCQ	63.0	66.9
Infer	31.9	61.3	Info-Cont	30.4	67.1	Open	37.0	59.5
Interpret	23.7	60.3	Info-Non-cont	34.1	67.4			
Evaluate	13.3	63.5						
Total	100.0	64.2	Total	100.0	64.2	Total	100.0	64.2

Since three of the booklets from 1998 are being administered again in 2004 to monitor trends since 1998, the breakdown of item characteristics for items from all five booklets is shown in Table 12.

Table 12 Breakdown of items by process, text type, and item type, combined 1998 and 2004 item pool (N=284; 135 new items and 149 retained items)

Process	Frequency	Percent	Text Type	Frequency	Percent	Item type	Frequency	Percent
Retrieve	105	37.0	Literary	107	37.7	MCQ	234	82.4
Infer	94	33.1	Info-Cont	99	34.9	Open	50	17.6
Interpret	66	23.2	Info-Non-cont	78	27.5			
Evaluate	19	6.7						
Total	284	100.0	Total	284	100.0	Total	284	100.0

Fit of final item pool to the framework

Table 13 shows the desired breakdown of items at each class level as specified in the framework, compared with the actual breakdown. At first class level the process match is good, but with a slightly higher proportion of interpret items and slightly fewer inference items than desired. The text type match is also good, but with a slightly higher proportion of informational-non-continuous items, and slightly fewer informational-continuous items than desired. At fifth class, the process match is also quite good, apart from the fact that there are fewer evaluate items, and more retrieve items, than desired. Also there is a slightly higher proportion of informational-

continuous, and a slightly lower proportion of informational-non-continuous items in the final item pool compared with the desired proportion. Part of the reason for this is the inclusion of one new informational-non-continuous passage at the end of one of the TARA 93/98 booklets in place of poetry items which were never used in scaling or reporting. Another reason is that at fifth class level, to measure trends in reading achievement, NAER 2004 must retain some of the TARA 93/98 booklets, so a very close match to the framework will not be possible until TARA 93/98 booklets are dropped in future NAERs. However, for NAER 2004, the slight imbalances are of no concern given that performance across the different items can be weighted, and also given the manner in which it is proposed to scale and report performance (see *reporting scales,* below).

Table 13 Desired and actual breakdown of items for NAER 2004: first and fifth classes

	First Class		Fifth Class		
Aspect	*Desired %*	*Actual %*	*Desired %*	*Actual %: New items*	*Actual %: All items*
Process					
Retrieve	45	46.3	30	31.1	37.0
Infer	35	30.5	30	31.9	33.1
Interpret	20	23.2	20	23.7	23.2
Evaluate	0	0	20	13.3	6.7
Text Type					
Literary	55	55.8	40	34.1	37.7
Info-Cont	25	21.1	30	30.4	34.9
Info-Non-cont	20	23.2	30	35.6	27.5
Item Type					
Multiple choice	100	100	67	63.0	82.4
Constructed response	0	0	33	37.0	17.6

Note. The desired percent correct is the midpoint of the range of values in Table 2.

Reporting Scales

It is proposed to scale results at both class levels to have an overall mean of 250 and standard deviation of 50 (to maintain comparability with performance at fifth class level in NAER 1998), and to report an overall combined reading scale score at each class level. It is also proposed to report results by subscales relating to text type and process underlying items, since some differences across subscales and groups of pupils may be apparent. For example, at first class, there are sufficient items to construct two text type scales: literary and informational, and two process scales: retrieval and inference/interpretation. There are more items at fifth class and it may be possible to report outcomes on three subscales for text type and three for process: literary, information-continuous, and information-non-continuous; and retrieval, inference, and interpretation/evaluation. Differences in achievement by item format will also be explored at fifth class level. The construction of these subscales depends on the fit of subsets of the items to the IRT model and on the premise that performance on subgroups of items are not too highly intercorrelated. Consideration will be given to developing categorical achievement ('proficiency') scales at each level, similar to those used in the reporting of the PISA 2000 results (e.g., OECD, 2001), which have numerical labels and qualitative descriptions of the likely success of pupils at each level in succeeding on items at the corresponding level.

CONTEXTS

Overview of Contexts

In addition to assessing pupils' reading achievement, NAER 2004 will gather information about children's experiences in various everyday contexts. Of particular interest for the purposes of NAER 2004 are the individual, home, school/class, and community contexts. This section presents an overview of the elements to be addressed within each context and the instruments that will be used to gather such information in NAER 2004.

Questionnaires to be administered are shown in Table 14. The elements within each of the contexts to be measured, and where this information is to be gathered, are shown in Table 15. It should be noted that, although described separately, the reciprocal influence that the contexts have on each other and on reading is recognised (i.e., home and school environments support each other, and the broader community context within which they are located influences the home and school). In addition to the questionnaires, the Department of Education and Science's Primary Schools Database will be used as a source of information. Both class levels receive the same instruments, with adaptations as appropriate to pupils' stage of development, instructional environment, etc. The Class Teacher, Learning-Support Teacher and Inspector Questionnaires were not administered in 1998 and will provide useful additional information about pupils classroom and Learning-Support environments. All questionnaires were updated since 1998. However, care has been taken to ensure that outcomes on the NAER 2004 questionnaires are comparable to the main outcomes on the NAER 1998 questionnaires.

The School Questionnaire used in NAER 1998 remains largely unchanged, except that some additional information on the first languages of pupils attending the school and the organisation of Learning-Support is sought. The Pupil Rating Form is also largely unchanged from 1998 except that it gathers more information on pupils' first language, and Learning-Support and resource teaching needs. The Pupil Questionnaire at fifth class retains most items from 1998 and includes a number of additional questions on pupils' attitudes to reading, motivation to read, and use of metacognitive strategies before, during and after reading. The Parent Questionnaire has been refined to include more detailed questions on socioeconomic status, parental attitudes to reading, and parental reading habits.

Table 14 Questionnaires to be administered in NAER 2004

Instrument	Respondent	Used in 1998
School Questionnaire	Principal	Yes
Class Teacher Questionnaire	Class Teacher	No
Learning-Support Teacher Questionnaire	Learning-Support Teacher	No
Pupil Rating Form	Class Teacher	Yes
Pupil Questionnaire	Pupil	Yes
Parent/Guardian Questionnaire	Parent/Guardian	Yes
Inspector Questionnaire	Primary Inspectors	No

Table 15 Elements to be addressed within each context in NAER 2004

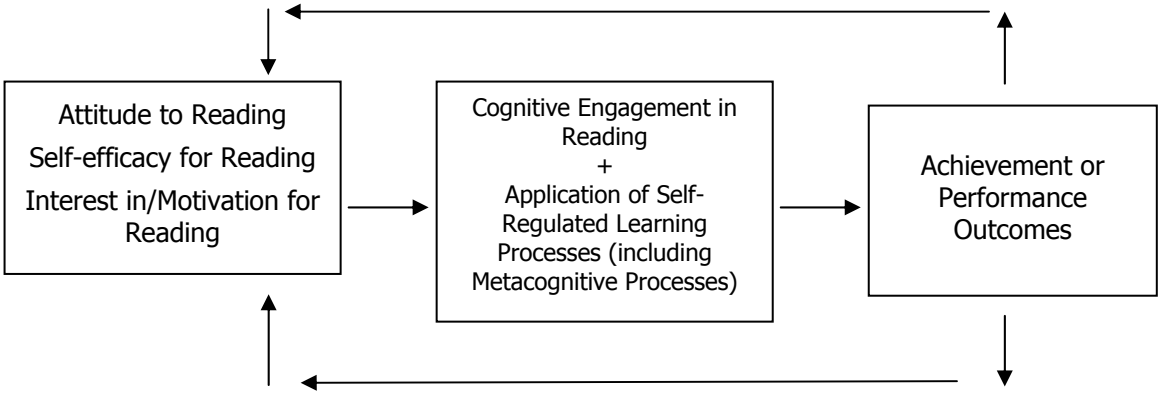
Context	Elements	Instrument
Individual	• Pupil characteristics (e.g., gender, age, first language) • Awareness and use of metacognitive (self-regulatory) strategies* • Attitudes towards reading; interest in and motivation towards reading • Self-efficacy for reading and learning • Engagement in leisure reading and reading preferences, use of libraries • Engagement in other leisure activities (e.g., playing computer games)* • Academic aspirations and goals* • Pupils' reading and writing standards (teacher assessed and self-assessed, and perceived teacher expectations) • Academic development (attendance of preschool, age began reading alone*) • Homework habits (e.g., frequency, time taken, type of homework task) • Pupils' engagement in class (e.g., behaviour, attendance, etc.) • Pupils' learning support status and involvement in learning support, resource and/or language support teaching **Assessed at fifth class only.*	Pupil Questionnaire Pupil Rating Form (teacher)
Home	• Family characteristics (e.g., family structure, language in the home, SES) • Home educational processes (e.g., child-parent literacy interactions) • Availability of academic guidance and support in the home • Parental reading habits and attitudes to reading • Disciplinary climate (e.g., rules around homework, and TV viewing) • Academic aspirations and expectations for children, standards set for children's achievement, and interest and knowledge in children's school life • Home literacy resources (e.g., access to books in the home)	Parent and Pupil Questionnaires
School / class	**School** (some of these variables are collected from the Department's Database) • School characteristics (e.g., size, gender composition, disadvantaged status, socioeconomic composition of school as indicated by mean SES of pupils attending, staff morale) • Pupil characteristics (e.g., ethnic / language mix of pupil body, attendance rates) • School material and personnel resources and facilities • Home-school links (Parents' Association, HSCL) • School management and organisation (e.g., school development planning, staff meetings, implementation of Primary School English Curriculum, assessment of reading, provision of learning support and language support) • Library resources and management/development of libraries • Perceptions of problems in the learning and teaching of reading • Implementation of *Learning-Support Guidelines* • Identifying/screening and grouping pupils for learning support **Class** • Teacher characteristics (e.g., gender, qualifications, teaching experience) • Class characteristics (e.g., class size, multi-grade classes) • Implementation of the Primary School English Curriculum • Participation of teachers in professional development in English • Instructional materials (e.g., usage of Reading Schemes and other materials) • Classroom resources and facilities (e.g., nature and use of library facilities, computers) • General instructional and assessment practices (e.g., grouping practices, homework assignments, times spent teaching English, lesson planning) • Reading instruction practices (e.g., frequency of employing various strategies outlined in the Primary School English Curriculum to promote engagement in reading, use of instructional strategies that promote awareness and use of metacognitive strategies)	School (principal) Questionnaire and Primary Schools Database Inspector Questionnaire Learning-Support Teacher Questionnaire Class Teacher Questionnaire
Community	• Local infrastructure (e.g., availability of library facilities, bookstores) • Urban/Rural location of school	School / Parent Questionnaire

Individual-Level Contexts

Pupils' attitudes, interests, behaviours, habits and learning strategies

In order to understand the performance of readers, and the reasons why they may choose to engage in or disengage from reading, it is necessary to consider associations between a range of attitudinal and motivational variables, use of cognitive strategies in reading, and reading achievement. Research in this area suggests that constructs such as interest in reading, attitude to reading, efficacy for reading, and motivation to read contribute directly to a student's decision to engage in reading, and, indirectly, to student achievement outcomes (Pintrich & De Groot, 1990; Schunk & Zimmerman, 1994) (Figure 6). Moreover, it has been shown that particular instructional environments and interventions can have positive and reciprocal effects on self-efficacy, motivation, use of self-regulatory learning strategies and achievement (Baker & Wigfield, 1999).

Figure 6 Hypothesised associations between interest, attitude, self-efficacy, motivation, engagement and reading outcomes



Attitude to reading

General attitude has been defined as "a learned disposition to respond in a consistently favourable or unfavourable manner, with respect to a given object" (Fishbein & Ajzen, 1975, p. 6). Alexander and Filler (1976) provide a reading-specific definition that is consonant with the view that reading attitude can be conceptualised along a continuum, with positive and negative extremes. They state that "attitude to reading is a system of feelings related to reading which causes the learner to approach or avoid a situation" (p. 1). According to McKenna, Kear and Ellsworth (1995), attitude to reading consists of three dimensions: prevailing feelings about reading, action readiness for reading, and evaluating beliefs about reading (efficacy for reading).

Both NAER 1993 and NAER 1998 (Cosgrove et al., 2000) looked at some dispositions associated with reading achievement. In 1998, students were asked to indicate their levels of agreement with five statements related to attitude: 'I like reading', 'Reading is boring', 'I wish books were shorter' 'It is important for me to do well at reading', and 'I read because it is a nice thing to do when I have time'. Pupils were generally well disposed towards reading, with almost 80% strongly agreeing or agreeing with the statement 'I like reading'. Correlations between the attitudinal items and performance on the overall TARA scale had absolute values ranging from .25 to .28.

In the recent OECD Programme for International Student Assessment (PISA), attitude to reading was measured using nine items with which participants (15-year olds) indicated their level of agreement. The statements included 'I read only if I have to', 'Reading is one of my favourite hobbies', 'I find it hard to finish books', 'I feel

happy if I receive a book as a present' and 'For me reading is a waste of time'. These items were combined to form a single measure of attitude to reading. In Ireland, the correlation between this variable and performance on a test of reading literacy was .42 (p. < .001) (Shiel et al., 2001). In a multi-level model of achievement in reading that included school- and student-level socio-economic status, effect sizes for students with poor, average and good attitudes to reading were -.3 SD, 0 SD and +.3 SD respectively, suggesting that attitude is an important variable in explaining reading achievement.

The number of items measuring attitude in NAER 2004 at fifth class level has been increased from 1998 using the nine PISA items and retaining two of the original NAER 98 items ('I like reading', and 'Reading is boring'), the latter two having provided the strongest correlations with reading achievement in NAER 1998. This selection of items consists of a mix of positive and negative statements about reading to which pupils must indicate their level of agreement. As they have already been trialled in previous studies, and assuming construct validity, we can be confident that this subset constitutes an accurate measure of attitude to reading. At first class level, just one question on attitude is included ('I like reading').

Parents/guardians of participating pupils are also asked about their attitudes to reading, making it possible, for the first time, to examine intergenerational differences in reading attitudes.

Self-efficacy for reading

Recent research on learning theory has shown that self-efficacy (self-beliefs) about factors such as one's learning capabilities, the content to be learned, and the setting in which the learning occurs are essential in motivating personal mastery of knowledge (Schunk & Zimmermann, 1994). In reading, children's mastery of text comprehension has been described as shifting gradually from social to self-regulated processes, as motivated by their growing sense of self-efficacy (Schunk & Zimmermann, 1997).

In PISA, self-efficacy for learning was measured as students indicated their level of agreement with four statements: 'I'm certain I can understand the most difficult material presented in readings', 'I'm confident that I can understand the most complex material presented by the teacher', 'I'm confident I can do an excellent job on assignments and tests' and 'I'm certain I can master the skills being taught.' In Ireland, a correlation of .24 was obtained between students' self-efficacy for learning and performance on combined reading literacy. It is quite possible that the correlation would have been stronger if the self-efficacy items focused more strongly on beliefs about reading, rather than on beliefs about learning in general. Self-efficacy in PISA also correlated moderately strongly with a range of other cognitive and motivational variables including control strategies (.55), elaboration strategies (.54) and control expectations (.67) (Cosgrove, 2003).

Given the high correlation between self-efficacy and motivation, it is perhaps not surprising that one of the dimensions in Wigfield's (1997) *Motivations for Reading Questionnaire* (MRQ) was entitled 'Reading Efficacy'. As well as a modified version of the one PISA item which was concerned with reading efficacy ('I can understand the most difficult books I am asked to read'), the current fifth class Pupil Questionnaire now includes adaptations of three items from the MRQ ('I know I will be a good reader next year', 'I am a good reader', and 'I learn more from reading than most other students in the class'). In his research, Wigfield found that the reading efficacy section, made up of four items, not only had good internal consistency, but positively predicted children's performance on standardised tests. We are therefore confident that the inclusion of this section will allow us to report a reliable measure of pupils' beliefs of self-efficacy in reading. At first class level, self-

efficacy for reading is tapped by just one item ('Do you think you are good at reading?'). Self-efficacy for writing is also tapped at first class.

Motivation for reading

According to Guthrie and Wigfield (2000), reading motivation is the individual's personal goals, values and beliefs with regard to the topics, processes and outcomes of reading. They view motivation as the link between frequent reading and reading achievement, and posit that highly motivated readers establish their own goals for reading, and provide opportunities for themselves to engage in extensive reading. According to Guthrie and Wigfield, motivation for reading can be viewed as being distinct from attitude, though in practice both are highly correlated with one another.

PISA also addressed students' motivations for learning, and included sets of items designed to measure general instrumental motivation (extrinsic motivation), effort and persistence (intrinsic motivation), co-operative learning, and competitive learning. None of these items was specific to reading, and, not surprisingly, correlations with PISA combined reading literacy were modest, ranging from .12 ($p < .001$) (effort and persistence) to .01 ($p = .225$) (co-operative learning). Therefore it is not planned to use the PISA motivation items.

Gambrell et al. (1996) developed a tool called *The Motivation to Read Profile* to assess motivation to read among students in Grades 2 to 6. The measure consists of a 20-item survey and a pupil interview schedule. In the survey, students are asked to indicate their responses to such statements as 'Reading books is something I like to do – never/not very often/sometimes/often', 'Reading is – very easy for me/kind of easy for me/kind of hard for me/very hard for me' and 'When I read out loud, I am a – poor reader/OK reader/good reader/very good reader'. Factor analysis confirmed that two constructs underpin the scale – students' self-perceived competence in reading (which is very similar to the self-efficacy construct discussed above), and their self-perceived performance relative to peers. Some evidence for the validity of the measure was evident in that significant differences in reading achievement were reported for students with high, average and low scores on the scale. It might be argued, however, that Gambrell et al.'s scale addresses the antecedents of motivation (e.g., self-efficacy or self-perceived competence) rather than motivation itself. Therefore it is not proposed to use these items in the Pupil Questionnaire either.

Finally, Wigfield (1997) developed a 54-item *(Revised) Motivation for Reading Questionnaire* designed for use with pupils in the senior primary grades. Again, using factor analysis, a number of dimensions of motivation were identified, including the following:

- Reading efficacy (e.g., 'I am a good reader')
- Challenge (e.g., 'I like to read hard, challenging books')
- Curiosity (e.g., 'I like to read about new things')
- Involvement (e.g., 'I read a lot of adventure stories')
- Importance (e.g., 'It is very important to me to be a good reader')
- Recognition (e.g., 'My friends sometimes tell me I am a good reader')
- Grades (e.g., 'I read to improve my grades')
- Reading for social purposes (e.g., I sometimes read to my parents')
- Social (e.g., 'I talk to my friends about what I am reading')
- Competition (e.g., I try to get more answers right than my friends')
- Compliance (e.g., 'I always do my reading exactly as the teacher wants it')
- Reading Work Avoidance (e.g., 'Complicated stories are no fun to read')

When these dimensions of reading were correlated with the reading frequencies of students, social, reading efficacy (as mentioned above), curiosity, involvement,

recognition, grades and reading importance had the strongest correlations with frequency. Wigfield concluded that children are more likely to read frequently when they feel efficacious about their reading skills, and are intrinsically motivated to read. The dimensions of motivation that most strongly predicted performance on a standardised test of reading achievement were reading work avoidance, social reasons for reading, reading efficacy and recognition. The correlations involving work avoidance and social reasons were negative. These outcomes confirm the multidimensionality of reading motivation, and the relative importance of intrinsic and extrinsic motivations.

Due to the high number of Motivation scales it was necessary to select a reasonable number of items whilst maintaining the validity of the measure. Wigfield's *Motivation for Reading Questionnaire* (MRQ) was selected as a template due to its reliability and the breadth of dimensions covered within motivation. From the 54 items making up 11 dimensions, 15 items were selected to form a general measure of motivation for reading for fifth class pupils, covering the following sections: challenge, curiosity, recognition, social reasons for reading, and reading work avoidance[3]. The items in the final pool were chosen because they were not academically oriented, since a measure of children's motivation for reading outside school instruction was desired. The pool was further condensed by choosing one item from a group of similar statements to represent that particular aspect of motivation. The final items were taken from the blocks Wigfield had found to be the best predictors of performance.

At first class level, one item taps each of these constructs via agree-disagree statements. The reading work avoidance construct is tapped indirectly in an item that asks pupils what they do when they come to a new word (responses are skip the word/try to sound it out/use the words around it to figure it out/ask someone for help).

Interest in reading

Interest has been defined by Hidi and Harkackiewicz (2000) as "an interactive relation between an individual and certain aspects of his/her environment" (p. 152). Hidi and Harkackiewicz distinguish between individual interest, which, by definition, is personal, and situational interest, which is generated by certain conditions in the environment. While some researchers use the terms 'interest' and 'intrinsic motivation' interchangeably, others (e.g. Schiefele, 1999) have argued that individual interest is an antecedent to cognition and determines the strength of an individual's intrinsic or extrinsic motivation to act in a particular situation. Thus, individual interest is viewed as a precondition of intrinsic motivation, with strong, stable and relatively well-focused individual interest functioning as powerful determinant of attention, memory and motivation.

In PISA, interest in reading was assessed using three items to which students had to indicate their level of agreement: 'Because reading is fun, I wouldn't want to give it up', 'I read in my spare time' and 'When I read, I sometimes get totally absorbed'. A composite 'interest in reading' variable consisting of students' aggregated responses across the three items was moderately associated with combined reading literacy ($r = .36$, $p < .001$) and strongly correlated with both frequency of reading ($r = .69$, $p < .001$) and attitude to reading ($r = .79$, $p < .001$) (Cosgrove, 2003).

Given the increased number of items relating to pupils' dispositions towards reading it was decided not to develop items that would constitute a separate Interest scale. This decision was made for two reasons. Firstly, there was a limited number of items which purported to solely measure interest in reading, with these items coming from PISA. With so few items it would be difficult to claim a reliable measure of Interest. Secondly, items similar to the Interest ones from PISA have been included in the

[3] The reading efficacy section of the MRQ has been used in the self-efficacy section referred to above.

questionnaire measuring associated constructs. 'Because reading is fun, I wouldn't want to give it up' and 'I read in my spare time', were both questions we felt had been covered in the Attitude section through 'Reading is one of my favourite hobbies', and through 'For me reading is a waste of time' (reverse scored). The third Interest item 'When I read I sometimes get totally absorbed', was also presented in a similar format in the new Motivation section of the questionnaire (see above). Consequently, so that the questionnaire could be kept concise while covering all the relevant dimensions, the items specifically for Interest have not been included, as similar and overlapping information will be placed elsewhere in the Pupil Questionnaire.

Frequency of reading

While the above measures relate to internalised constructs, the Pupil Questionnaire also collects information on pupils' actual reading behaviour and the frequency with which they engage in various types of reading (stories/novels, information texts, magazines, comics, newspapers, and emails/web pages at fifth class level, and books, magazines and comics at first class level). Leisure reading (outside school time) was chosen to maintain consistency with the motivation questions described above. This question is based on an item taken from the PISA 2000 Student Questionnaire and is more detailed than the questions on frequency of reading used in NAER 1998. The Pupil Questionnaire will also ask pupils how often they use public and school libraries (in NAER 1998, pupils were only asked about public library use).

Research has consistently demonstrated a positive correlation between reading achievement and time spent reading for fun (Campbell et al., 2001). For example, Cosgrove and Morgan (2000), in secondary analyses of NAER 1998 data, reported a correlation of .31 between the frequency of reading books in leisure time and overall reading achievement. Book-reading was a better predictor of achievement than newspaper reading ($r = .09$) or magazine/comic reading ($r = .04$). Furthermore, the scores of pupils who read books daily or nearly every day (44% of pupils) were, on average, 46 points, or nearly one standard deviation higher than the mean scores of those pupils who said that they hardly ever or never read books (7% of pupils). In PISA, students who reported never reading (33% of students) had a mean reading score that was 67 points or two-thirds of a standard deviation lower than students who read for half an hour or more per day (20% of students) (Shiel et al., 2001).

NAER 2004 will also gather information from parents/guardians for the first time on the frequency with which they engage in various types of reading activities, allowing a comparison of parental and pupil reading activities.

Reading preferences

In NAER 1998 two items on reading preferences were subsumed under the general category of attitude to reading in the fifth class Pupil Questionnaire, and it was felt that a separate measure of reading preferences should be included in NAER 2004 as distinct from attitude. Three items from the MRQ were selected to assess reading preferences (preference for adventure, fantasy and mystery texts). Items tapping reading preferences were not included in the first class Pupil Questionnaire.

Metacognitive processes

Transcending the processes of comprehension are the metacognitive processes and strategies that allow pupils to monitor their comprehension and adopt appropriate strategies as they read (Campbell et al., 2001). Metacognitive knowledge includes an awareness of self as a learner, understanding of what the task demands, and motivation to use comprehension strategies (Shiel & Murphy, 2000). Strong metacognitive knowledge and an awareness of metacognitive strategies have been shown to be characteristics of good comprehenders (Schmitt, 1990; Shiel & Murphy, 2000), while instruction in reading comprehension strategies is believed to lead to improvements in reading comprehension, particularly among low-achieving readers

(National Reading Panel, 2000). Research suggests that if less able readers are not given formal instruction in cognitive strategies to enhance their comprehension, they are unlikely to learn, develop and use these strategies (National Reading Panel, 2000). Reading comprehension strategy instruction appears to be most effective when multiple strategies are taught, and where the strategies are explicitly explained; for example, where the teacher describes the strategies and demonstrates, models and guides the reader in their use. After learning and applying these strategies with the teacher's assistance, the reader should gradually internalise and begin to use such strategies independently in the course of their everyday reading (National Reading Panel, 2000).

A range of reading comprehension strategies have been highlighted in recent literature on reading, and include previewing, purpose setting, predicting and verifying, generating questions and interpretations, using background knowledge, self-monitoring and self-questioning, applying fix-up strategies, and summarising content.

The Primary School English curriculum makes reference to some comprehension strategies that children should be encouraged to use/apply. Such strategies include predicting, checking, confirming, and self-correcting. The curriculum suggests that first/second class pupils should be enabled to:

- Self-correct reading errors when what is read does not make sense
- Adapt his/her reading style for different purposes (implying that the child should be aware of purpose for reading and adjust approach based on this)
- Adopt an active approach to text by posing one's own questions.

There are fewer references to such strategies at fifth/sixth class level (possibly because pupils at this stage will have already acquired a range of strategies), although the Primary School English Curriculum suggests that pupils at this level should be able to develop study skills such as skimming, scanning, note-taking and summarising.

Drawing on the *Metacomprehension Strategy Index* (MSI) (Schmitt, 1990) we have created a sample of items to test fifth class pupils' use and awareness of the various metacognitive strategies outlined above. The MSI is split in to three sections, covering possible strategies used before, during and after reading that increase understanding. For each item, four statements about what the reader may do to help their understanding of a story at each stage of reading is presented, and the respondent is required to choose the statement most similar to their behaviour or approach. These items have been significantly modified in order to fit more with the purpose and format of NAER 2004. The context was broadened to refer to any text the pupil may read, not just stories, as understanding of expository and document texts are also of interest. Items were selected from each stage of the reading process (before, during and after), those inappropriate to fifth class level were excluded and the format altered so that response options became the *frequency* of carrying out the strategy of interest. The final selection consisted of five items referring to strategies used before reading, four items relevant to strategies used during reading and four after reading.

Some items assessing metacognitive strategies used by first class pupils (e.g., 'I read difficult parts of a story again to make sure I understand it') were piloted in May 2003 but the response pattern suggested that pupils did not comprehend the questions and it was decided to drop these items, which were felt to be developmentally inappropriate.

Reading development

The two most recent national assessments of reading in fifth class included an item designed to provide retrospective data on pupils' reading development ('At what age did you begin to read books on your own?'). In NAER 2004, it is planned to retain those items. This will enable a comparison of pupils' reading development over the years (for example, to ascertain whether the age at which pupils began reading independently has changed). This question is not included at first class level.

Homework practices

Although one might expect frequent involvement in, and time spent on, homework to be associated with higher levels of reading achievement, the evidence from research into the relationship between homework and achievement at primary school level is inconclusive (Sharp, Keys & Benefield, 2001). In NAER 1998, a weak significant negative relationship between time spent on homework and reading achievement was found (Cosgrove et al., 2000). However, Cosgrove et al. (2000) note that the relationship between homework and achievement may not be accurately reflected in an analysis of homework and achievement in isolation, since the amount of time spent on homework is influenced by other factors such as pupils' ability (i.e., less able pupils may take longer to do homework; more able pupils may elect to do extra homework). There is also limited evidence regarding the impact of different types of homework assignments on achievement (Sharp et al., 2001). For example, Cosgrove et al. (2000) found a positive correlation between achievement of fifth class pupils and both frequency of reading a library book for homework, and frequency of answering comprehension questions for homework. However, there was no relationship between achievement and frequency of learning spellings for homework.

The two most recent national assessments of reading in fifth class included a number of items designed to assess pupils' homework practices. In NAER 2004, these items will be retained, with the addition of new items requesting further detail on the nature of the homework tasks pertaining to English. At first class level, pupils are asked how often they get homework, and do various types of homework (e.g., learning spellings, doing silent reading).

Academic aspirations and goals

The two most recent national assessments of reading in fifth class included a number of items designed to assess pupils' attitudes to school (liking school) and their expectations about how long they will spend at school. In NAER 2004, it is planned to retain those items at fifth class level. This information can be compared with data gathered in NAER 93/98, and can be linked to parents' and teachers' responses to items about academic expectations in the Parent and Teacher Questionnaires. In NAER 1998, Cosgrove et al. (2000) found significant positive associations between pupils' liking of school and their expected level of attainment, and reading achievement. New items in NAER 2004 at both first and fifth class levels include pupils' perceptions of their teacher's expectations of them, which will provide an interesting comparison with both their own self-efficacy ratings and their performance on the reading test.

Pupils' engagement in school

At both class levels, the pupils are asked whether they like school or not. In addition to this attitudinal measure, behavioural measures of engagement will be gathered through the Pupil Rating Form; e.g., teachers' ratings of pupils' behaviour and participation in class, and information on their attendance rates. These measures are of interest given that several research studies have demonstrated an association between poor attendance and low attainment (e.g., Mullis et al., 2001). Cosgrove et al. (2000) also reported a small but significant positive association between reading achievement and attendance ($r = .19$) in NAER 1998. The small magnitude of the correlation may be partly because there was little variation in attendance rates.

Teachers' ratings of pupils' reading standards and learning needs

Reading standards

In 1998, teachers rated pupils' overall reading ability in terms of both a qualitative scale (ranging from advanced to weak) and in terms of the class level they would place them at (ranging from post-primary to third class or lower). About 10% of pupils were placed in the lowest ability categories and the correlations between these ratings and pupils' reading achievement were strong (r = .68 for qualitative ratings and r = .66 for class level ratings). However, the interpretation of these items is likely to vary across classes and schools since the standards against which teachers compare pupils vary. Therefore, these items have been refined to refer explicitly to *national* standards, and the qualitative rating item expanded to include ratings on different aspects of reading and writing. A new item asking teachers to place pupil performance in a percentile category based on outcomes on a standardised reading test is also included. Teachers are also asked (as they were in 1998) to indicate how they expect pupils to perform on reading tasks in the future; e.g., teachers of fifth class pupils are asked to rate how well they expect pupils to cope with the reading and writing tasks of post-primary level. In 1998, about 2-4% of pupils were judged by teachers as not being able to cope with these tasks, and a further 12-23% rated as needing assistance with these tasks.

Learning needs

Teachers of pupils will be asked to indicate, for each pupil, whether they are in receipt of Learning-Support teaching, resource teaching, and/or language support teaching, since educational needs of pupils can be expected to impact on their achievement. Teachers will also be asked whether any pupils not in receipt of additional teaching support outside regular class teaching should, in their opinion, receive it. These indicators can serve as a measure of the adequacy of Learning-Support resources (from the perspective of teachers) as well as an individual pupil characteristic.

Home and School/Class Contexts

Both the home and the school/class contexts are important in children's early reading experience. Primary school children develop the skills and attitudes associated with reading, both formally, through structured instruction in school, and also informally, in daily activities at home and school (Campbell et al., 2001).

Home environment

The Primary School Curriculum recognises the role of the parents[4] as the child's primary educators, and the powerful influence of the home environment on children's acquisition of behaviours, attitudes and skills associated with reading. The role of parents in the early acquisition of reading is critical (Campbell et al., 2001).

An initial appreciation of the usefulness and purposes for reading can be fostered when parents read aloud to young children. Other activities such as drawing, play with books and other printed materials, and writing (e.g., writing names and forming letters) also promote literacy and help children develop their awareness of text. The development of positive attitudes and an enjoyment of reading will also be fostered through early association with printed text in the home. According to Campbell et al. (2001) and Kellaghan, Sloane, Alvarez & Bloom (1993), home factors associated with reading include:

- Language in the home—A knowledge of language is critical for the acquisition of reading. Children whose language development is below that expected of children at their age level, or whose home language and dialects are different

[4] The term 'parents' in this document is used to refer to both parents and guardians.

from those used at school, may be at an initial disadvantage when learning to read.

- Home literacy resources—A print-rich environment in which children have ready access to reading material and resources has also been shown to be positively associated with the literacy achievement (e.g., Martin & Morgan, 1994). In addition to providing children with an opportunity to practice and learn skills associated with literacy, the availability of such resources in the home can influence children's attitudes, as it conveys that learning to read is useful and valuable.
- Home educational processes or activities fostering reading—Reading achievement has been shown to be higher among pupils whose parents take an active role in their children's learning. For example, research has consistently linked the reading of books to children at home to early literacy and school success (Dickinson & Tabor, 2001). Parents who frequently engage in different literacy activities with their children can foster positive attitudes towards reading, as well as provide modelling and guidance which promotes the development of effective literary practices. Overall, Dickinson and Tabor (2001) found that home support for literacy (i.e., the quantity of books owned, the frequency of reading, and variety of reading activities) was predictive of children's early literacy skills.
- Availability of academic guidance and support, home-school links, and academic aspirations and goals—Research has shown that discussing schoolwork and reading experiences at home bears a positive relation to reading achievement (e.g., Kellaghan et al., 1993). Parents who hold high academic expectations for their children and demonstrate positive attitudes towards school tend to reinforce the value of learning to read, monitor their children's completion of schoolwork, and provide encouragement for their children.
- Disciplinary climate of the home—Living in a structured environment where there are rules and a routine is an important for children's development since through this structure, children can learn to prioritise tasks and develop good habits around organising and carrying out activities. Thus the Parent Questionnaire asks about the making of rules regarding the watching of TV, videos/DVDs, and playing computer games.

Socio-economic status

Socio-economic status (as indicated by, e.g. occupation; level of parental education) has been shown consistently to be related to literacy achievement (e.g., Kellaghan, Sloane, Alvarez & Bloom, 1993). However, the magnitude of this relationship varies across different cultures (OECD, 2001) and according to which measure of socio-economic status is used. In NAER 1998, SES was measured according to parental education and medical card status and it was found that pupils whose parents had a medical card scored 25 points (half a standard deviation) lower than those whose parents did not. The relationship between education level and achievement was moderate at .35. In PISA, parental education was also related to achievement in reading ($r = .21$). PISA also gathered information on parental occupations coded according to the International Standard Classification of Occupations (ISCO) (Ganzeboom & Treiman, 1996), which places occupations along a scale ranging from 0 to 90. The correlation between parental occupation and reading achievement in PISA was .31. The existing socio-economic measures are retained from NAER 1998, and questions about parental occupation (coded according to ISCO) and employment status have been added.

Family structure and size

Over the past decade, the overall number of one-parent households has increased by about 20% (NESF, 2001). One-parent households appear to be particularly vulnerable to poverty and disadvantage in Ireland compared to those in some other

EU countries, and that this trend has increased over recent years. For example, the NESF (2001) has noted that 73,000 one-parent households are in receipt of social welfare. This represents a six-fold increase over the past two decades. Recent surveys of educational achievement suggest that about one in eight children are living in one-parent households. For example, in PISA 2000, approximately 12.8% of students were living in one-parent households. In NAER 1998, around 12.0% of pupils were living in one-parent households. On average, children in one-parent households have lower reading achievement than children in two-parent households. For example, in NAER 1998, it was found that only 7.4% of pupils living in one-parent, female-headed households scored more than one standard deviation above the mean, compared with 18.2% of pupils living with both parents (Cosgrove et al., 2000). Family size was also negatively related to achievement in NAER 1998 (r = -.19). Therefore, NAER 2004, as with NAER 1998, will collect information on both family structure and family size, as indicated by the number of siblings and others living in the household.

School/class

The school is typically the main location where formal learning and educational activities take place. The school context therefore plays a key role in children's early acquisition of skills and attitudes associated with reading literacy. Factors in the school which may affect the learning of reading include:

- School environment and resources—A stable learning environment, where there are few discipline problems and where pupils feel a sense of security is important. As well, pupils' reading development may be enhanced in environments where staff demonstrate positive attitudes towards pupils and learning, and where staff collaborate on activities that foster reading (Postlethwaite & Ross, 1992; Sammons, Thomas, & Mortimore, 1997; Campbell et al., 2001). The two most recent national assessments of reading in fifth class included a number of items designed to assess school management and organisation (e.g., nature and content of school policy, staff meetings, school plans, assessment policy). These items will be retained in the NAER 2004 School Questionnaire. Furthermore, items designed to elicit information on the implementation of the Primary School English Curriculum will be included in the 2004 School Questionnaires, as will items asking principals' opinions on the usefulness of the Curriculum in school planning. More detailed information on the Primary School English Curriculum, particularly with regard to the implementation of approaches to the teaching of reading, will also be sought in the Teacher Questionnaire. The School and Teacher Questionnaires also ask about perceived problems (if any) in reading instruction at both the classroom and school level.

- School and classroom resources: The availability of resources in the school can also have an impact on pupils' learning. Adequate classroom space and facilities, accessibility of learning materials, and availability of trained and qualified staff, including a Learning-Support teacher, can help strengthen the English curriculum in the school (Campbell et al., 2001).

  The availability of a library corner (for junior pupils) or classroom library (for senior pupils) can promote positive attitudes towards reading. Furthermore, research has indicated that exposure to a variety of texts and text types is associated with reading achievement. In fact, the English curriculum notes that "the importance of the classroom library in enriching the child's reading experience cannot be over-emphasized" (p. 53). It suggests that children should play a role in the organisation of the classroom library, and that its layout should be similar to that of a public library. Furthermore, it suggests that the classroom library should cater for all levels of interest and ability. The

NAER 1998 School Questionnaire included a series of items designed to gather information on the library and computer facilities available in schools. Items in the 2004 School Questionnaire have been expanded to ascertain more information about the types of library facilities in each school, staff responsible for the libraries and use and availability of computers. The Teacher Questionnaire will include items that will provide information of a similar nature but specifically within the classrooms. Furthermore, to get an idea of the types of texts used in reading instruction items asking teachers to indicate their usage of Reading Schemes and other reading materials will be included in the Teacher Questionnaire.

The items in NAER 1998 regarding the school staff and availability of remedial teaching have been modified and updated. Information is requested about general staff, Learning-Support requirements and teaching in the school as well as resource teaching for pupils with specific learning disabilities. In addition to retaining these items in the School Questionnaire, the Learning-Support Teacher Questionnaire will be of use in assessing the type of Learning-Support resources that are available and how such resources are structured/allocated.

- Teacher training and preparation: The Primary School English Curriculum notes that "it is the quality of teaching more than anything else that determines the success of a child's learning and development in school" (p. 20). Since the extent of a teachers' training in teaching reading may have an impact on pupils' acquisition of reading literacy (Campbell et al., 2001), principals in the school are asked to indicate the number of teachers who do and do not have a basic teaching qualification. Two items will also be included to assess the amount and quality of professional development available to teachers in that school with in the context of the Primary School English Curriculum. Learning-Support and classroom teachers are also asked about their experience and qualifications, and the availability and perceived quality of in-career development courses.

  Much research has also been conducted with a view to identifying those practices associated with effective teaching of reading (e.g., Hoffman, 1986; Wray et al., 2001; Taylor & Pearson, 2002). For example, using a variety of approaches for teaching reading, and tailoring reading instruction to the needs of individual pupils has been identified as effective. Thus, it is planned to include items designed to gather information on teachers' general instructional and assessment practices, as well as their reading instruction practices, in the Teacher Questionnaires. Given that teachers' attitudes and expectations for their pupils may also have an impact on pupils' learning and success in school (Lumsden, 1997), it is proposed that items assessing teachers' expectations for their pupils' future academic success be included in the Teacher Questionnaire in addition to those items measuring teacher expectations with respect to individual pupils in the Pupil Questionnaire.

- Classroom environment and structure: Children's experience in the classroom is particularly salient, since they spend many hours per day in this environment. Class size and structure can have an impact on pupils' reading development. For example, some research suggests that children may benefit from smaller class sizes in the early grades (Campbell et al., 2001). As well, class structure can influence children's reading development. For example, a classroom environment which promotes interaction and provides opportunities for informal and formal discussions about reading among pupils may promote reading development (Campbell et al., 2001). Thus, information regarding class size

and structure (e.g., whether multi-grade or single-grade) will be gathered via the Teacher Questionnaire.

- Home-school links: Ideally homes and schools should play complementary, mutually reinforcing roles in children's education (Kellaghan, Sloane, Alvarez & Bloom, 1993). The home plays a critical role in the early development of skills and attitudes (e.g., language skills, interpersonal skills, motivation to learn) that are related to the work of the school and children's school learning. Research has shown that the development of home-school links can have a significant positive impact on children's cognitive development, school performance, and social functioning (Kellaghan et al., 1993). The items on home-school links in NAER 1998 have been further developed in NAER 2004 so that more information can be gathered about programmes devoted to encouraging children to read outside school. Parental interest/attendance will also be investigated, as well as Parents' Association groups and activities. These items should provide information indicative of the extent of home-school links, which can then be correlated with pupils' performance on the reading tests.

In addition to obtaining information on the school factors outlined above, information on school characteristics including school size, gender composition, designated disadvantage status, and socioeconomic composition will be gathered in NAER 2004 to allow for an investigation into the effects of the various aspects of school resources and circumstances on children's performance in reading.

Community Contexts

The broader environment in which children live will also have an impact on children's learning experience. Communities vary in terms of their resources and organisational features, and these factors will likely have an impact on homes and schools, and ultimately on children's experience in learning to read (Campbell et al., 2001). For example, in a comparison of the ten highest and ten lowest scoring countries participating in the IEA Study of Reading Literacy, access to books in the community was shown to be related to mean achievement at country level (although when wealth was taken into account, the importance of access to books diminished some) (Elley, 1992). It is proposed that items be included in the NAER 2004 school and Parent Questionnaires with a view to gathering information on the availability of literacy resources (e.g., public libraries) in the community. It should be noted, however, that the NAER 2004 framework is not designed to provide in-depth information about pupils' broader community environments.

TECHNICAL INFORMATION

Sample Design

Overview

The design is a two-stage stratified cluster sample, whereby schools are selected at the first stage and pupils selected at the second stage. Schools are sampled with probability proportional to size, and at the second stage of sampling, intact classes of pupils are selected. All classes in small (<21 pupils at the class level) and medium (21-34 pupils) schools are selected, while two classes are selected with equal probability from large (>35 pupils) schools. This procedure differs slightly to that of NAER 1998, where a random sample of a fixed number of pupils was sampled at the second stage of sampling. However, the number of schools and pupils selected is intended to ensure sufficient sampling precision to produce reliable, comparable population estimates (Cochran, 1977; Kish, 1957), and analysis methods will take the clustered nature of the sample design into account. Because testing takes place at first and fifth classes, and also because a survey of mathematics achievement is being carried out concurrently at fourth class level, there are a total of seven strata, based on the three class levels and enrolment size:

1. Schools with pupils selected from first class only (n = 16)
2. Schools with pupils selected from fourth and fifth classes only (n = 36; 10 small schools, 10 medium schools, and 16 large schools)
3. Schools with pupils selected from first, fourth and fifth classes (n = 100; 24 small schools, 24 medium schools, and 48 large schools).

Target population

The target population consists of all pupils in first and fifth classes in primary schools in Ireland in May 2004. All pupils in mainstream (ordinary) classes in primary schools are eligible to participate. However, pupils attending private schools (1.04% of first class pupils and 1.24% of fifth class pupils), special schools (0.54% of first class pupils and 0.98% of fifth class pupils), or special classes in ordinary schools (1.77% of first class pupils and 2.14%) are excluded[5]. Hence the defined target population includes, or covers, approximately 96.65% of all pupils in first class and 95.64% of all pupils in fifth class in the country.

Schools were excluded either because they were listed as not having any pupils at two of the three[6] grade levels (20 of 3155 schools[7]) or, in the first class only stratum, they were listed as having fewer than five first class pupils (59 of 3155 schools). These exclusions cover just 2.5% of the total number of schools, and less than half a percent of pupils at each class level (0.26% at first class, and 0.37% of fifth class).

Table 16 shows the number of primary schools (N = 3,155) with at least one pupil at each class level, and Table 17 shows the number of schools with pupils at the three class levels.

[5] These estimates are extrapolated from the *Annual Statistical Report* of the Department of Education and Science (2000-2001, p. 15).

[6] Three grade levels rather than two are mentioned here since the sample design incorporates the sample of fourth class pupils for the concurrent national assessment of mathematics.

[7] After sampling it transpired that 12 schools on the sampling frame had amalgamated into six schools at the end of 2003, that a further six schools closed at the end of 2003. One sampled school had amalgamated with another and was replaced with the amalgamated school. Another sampled school was located in a remote area and was replaced with its first replacement school. In addition, eight new schools with provisional recognition were established at the beginning of the 2003-2004 school year. No information is available on the enrolment of these new schools and they were not included on the sampling frame.

Table 16 Numbers of schools which have at least one pupil at each class level

Classes category	Does school have at least one pupil at this class level?			n Schools	% Schools
	First	*Fourth*	*Fifth*		
1st, 4th & 5th	Yes	Yes	Yes	2,707	85.8
1st only*	Yes	Yes	No	32	1.0
1st only*	Yes	No	Yes	36	1.2
1st only	Yes	No	No	119	3.8
4th & 5th only	No	Yes	Yes	241	7.6
Excluded	No	Yes	No	3	0.1
Excluded	No	No	Yes	10	0.3
Excluded	No	No	No	7	0.2
				3,155	100.0

*Since the total number of pupils in 5th class is used as the measure of size, the fact that these two categories are missing 4th and 5th class pupils presents a problem. Rather than exclude these, they have been included in the '1st class only' category. A number of the schools which have 1st and 4th class pupils are junior schools with sizeable numbers in 1st through 4th. The schools with 1st and 5th pupils but no 4th class are mostly small schools. 4th and 5th class pupils in these '1st only' schools are excluded regardless of whether or not the school has 5 or more pupils in first class or whether or not it is selected.

Table 17 Number of schools with pupils in the different class levels (total number of schools=3,155)

Class level	n schools
1st class	2,894
4th class	2,983
5th class	2,994

Selection of schools

In the sampling frame of schools eligible for selection, schools were divided into strata based on enrolment size: small schools (fewer than 21 pupils enrolled at the class level), medium schools (between 21 and 34 pupils enrolled at the class level); and large schools (35 or more pupils enrolled at the class level). At first class level, a fourth stratum comprising schools with pupils in first class only, was identified. At fifth class level, six strata were identified: small, medium, and large schools, with pupils enrolled at first and fifth class; and small, medium, and large schools, with pupils enrolled at fifth class, but not first class. The stratification of schools is more complex than that for NAER 1998, since it entails selecting schools for participation in three surveys. This was done to minimise the number of schools cost and to maximise the efficiency of the survey design. Within these strata, schools are sorted by designated disadvantaged status, area/language of instruction (i.e., Gaeltacht, all Irish, or English language), proportion of female pupils, and measure of size. This sorting process ensures that a representative mix of school types is sampled.

The number of schools and pupils to be selected depends in part on the amount of clustering occurring within schools (Kish, 1957). For example, if differences between schools' achievement is large, then the design effect will also be large, resulting in a smaller effective sample size. The minimum effective sample size for surveys of educational achievement is usually cited as 400; that is, after controlling for the clustering of pupils within schools, the sample selected should be equivalent to a simple random sample of at least 400 pupils (Cochran, 1977). In Ireland, differences between schools at primary level are relatively small. For example, in NAER 1998, the proportion of variation in achievement between schools was about .13 (the so-

called intra-class correlation or rho). At first class, there is some evidence that rho is a little higher than at fifth. For example, the rho for the sample of first class pupils that took part in the standardisation study of the Drumcondra Sentence Reading Test (DSRT) is 17. The rho associated with the sample of first class pupils that participated in the standardisation study of the Drumcondra Primary Reading Test (DPRT) is .28. Another factor affecting the number of pupils and schools to be selected is the cluster size, i.e., the number of pupils to be selected in each school. An average cluster size of at least 25 is recommended if one is using multilevel analysis techniques where stable school-level estimates are required (Kish, 1957). A final factor to be considered in the development of the sample design is the number of pupil responses per test item required to scale the data using IRT – 900 to 1,000 responses per item is recommended (Mislevy & Bock, 1990). Taking these considerations into account, a total of 152 schools are sampled: 100 of these with pupils at both first and fifth class levels, 36 with pupils at fifth class only, and 16 with pupils at first class only (Table 18).

Table 18 Numbers of schools to be sampled overall

Classes sampled	**Small (<21)**	**Medium (21-34)**	**Large (35+)**	**Total**
First Class				
1st, 4th & 5th	26/2003	26/442	48/262	100/2707
1st only		16/128		16/128
Fifth Class				
4th & 5th only	10/99	10/42	16/100	36/241
1st, 4th & 5th	26/2003	26/442	48/262	100/2707
				152/3076

Note. Each cell shows the number of schools to be sampled out of the total number of schools in that stratum. Under- and over-sampling of strata can be redressed by using sample weights.

Selection of pupils

Unlike NAER 1998, where a random sample of pupils was selected across classes at fifth class level, intact classes will be sampled for NAER 2004. The reason for this is that the sampling of intact classes is administratively less complex which is desirable given that the assessment will occur at two class levels with concurrent administration of a national assessment of mathematics at fourth class; however, comparability with NAER 1998 at fifth class level will not be affected, given that the analysis methods used will produce standard errors associated with achievement estimates will take the clustered sample design into account. In small and medium stratum schools, just one class will be selected since it is likely that there is only one class per grade level, while in the large school stratum, two intact classes will be selected. Where there are more than two classes at a given grade level, two classes will be sampled at random.

Assuming an average of 10 pupils selected at each class level in small schools, 25 in medium schools, and 50 in large schools, it is estimated that a total of 4240 pupils will be sampled at first class level, and 4540 pupils will be sampled at fifth class level.

Exemption of pupils

Certain pupils may be exempted from the assessment and principals and teachers will receive detailed guidelines on exemptions. All pupils are encouraged to participate, but some may find the NAER 2004 assessment situation upsetting or frustrating. Thus, pupils enrolled in ordinary classes with moderate or severe learning disabilities, pupils with a physical disability that prevents them from participating, and pupils with less than one year's instruction in English, may be exempt from the assessment.

Multigrade classes
It may be noted that, in schools where there are multigrade classroom arrangements, all the pupils at the class level of interest in the multigrade class will be invited to participate, and Inspectors will take a more active role in these schools to assist the teacher with the administration of the assessment. Both Inspectors and teachers receive full information on the administration procedures to assure uniformity in assessment conditions across schools.

Schools with multigrade classes which include pupils at fourth and fifth class level are potentially problematic since assessments of reading and mathematics are concurrent, and the administration procedures for these differ. When schools are invited to participate, they will be asked for details of multigrade classroom arrangements. If a combined fourth/fifth class is in place in any of the schools, they will be consulted on a case by case basis. In some cases it may be feasible to administer both assessments. However, in other cases, it may not be administratively feasible to do so and in such cases, schools will be invited to participate in one or other survey.

Test and Questionnaire Analysis

In analyses, it is useful to distinguish between different types of variables. The outcome measure, reading achievement, will be reported in several ways:

- Measures of central tendency (mean, median)
- Measures of the dispersion or variability (standard deviation, scores at specific percentile points)
- An overall achievement scale and subscales based on each reading process/purpose
- It is proposed to develop a four- or five-level level categorical description of children's achievement, similar to the proficiency levels derived from the PISA 2000 data (OECD, 2001). However, the technical complexities and political issues associated with such a task need to be taken into account during this process (e.g., National Academy of Education, 1993), as well as the time available for this work.

Variables derived from the questionnaires take one of four basic forms:

- Numeric or count variables, e.g., the number of pupils in the school, or the number of library books in a classroom
- Categorical indicator variables, i.e., denoting membership to a group such as male/female, designated/not designated disadvantaged, etc.
- Categorical interval variables, i.e., points along an interval scale, such as a response to an attitudinal item ranging from 'strongly agree' to 'strongly disagree'
- Composite continuous variables. Some composite continuous variables are derived from numeric variables, e.g., the percentage increase in the number of library books in a school is computed from two variables, one indicating the total number of books and the other indicating the number of new books bought that year. Other composites are derived from responses to Likert type items such as the attitudinal example above. If a number of items measure a common construct, such as attitude to reading, then combining these responses gives a much more reliable measure of that construct than a single item; also, missing responses can be imputed or inferred from given responses. This type of composite has the added advantage that it can be transformed to any scale, such as mean of 5 and standard deviation of 1, or mean of 100 and standard deviation of 20. These composites will be constructed using regression-based estimation methods (e.g., principal components).

Continuous variables (counts, composites) can be simplified if the analysis requires. For example, pupils' scores on a metacognitive reading strategies scale can be split

into three, resulting in discrete groups of pupils – those with high, medium and low strategy use.

Outcomes on the achievement measure will be described initially, and then achievements of subgroups compared using bivariate analyses such as a t-test or chi-squares (to answer questions like: do boys perform significantly better than girls? Are girls more likely than boys to be in the bottom tenth percentile?). Other bivariate statistical tests such as a Pearson's correlation will be carried out to examine the relationships of background characteristics with achievement (to answer questions like: how strong is the relationship between pupils' socioeconomic background and achievement?).

Bivariate analyses have limitations, however since only one explanatory variable at a time is examined. Further, such analyses do not take into account the fact that pupils are clustered within schools. Multilevel modelling overcomes both of these limitations in that it allows one to examine the effects of multiple variables (and their interactions) simultaneously, at both the school level and the pupil level (Raudenbush, Bryk, Cheong, & Congdon, 2000). These models are useful since they can often reveal large differences between the unadjusted effect of a variable (e.g., the relationship *solely* between frequency of reading and achievement) and the adjusted effect (i.e. that relationship after adjusting for a host of other variables). They also provide information about whether an effect for a particular variable is constant across schools/classes, or whether is varies across schools. This technique was used to explore relationships between multiple school and student explanatory variables and 15-year olds' achievement in reading, mathematics and science in PISA 2000 (Shiel et al., 2000). It is also planned to develop a multilevel model of achievement at both first and fifth classes in analyses of the NAER 2004 data.

Finally, an examination of the variation in achievement will be carried out. The variation in achievement can be partitioned into three components: that which is between schools, that which is between classes, and that which is between pupils. This is of policy interest since the lower the between school variance, the more homogenous schools are with respect to achievement, and by inference, the more equitable (less selective) the school system is. It is also of interest to estimate the proportion of variance that can be attributed to classes and to individuals. Variance components will be examined at both first and fifth class levels.

REFERENCES

Adams, R.J., & Wu, M. (Eds.) (2002). *PISA 2000 technical report*. Paris: OECD.

Adams, R.J., & Cartensen, C. (2002). Scaling outcomes. In R.J. Adams & M. Wu (Eds.), *PISA 2000 technical report* (pp. 149-162). Paris: OECD.

Alexander, J. E., & Filler, R. C. (1976). *Attitudes and Reading.* Newark, DE: International Reading Association.

Assessment Systems Corporation (1995). *ITEMAN for Windows (Version 3.50).* St Paul, MN: Author.

Baker, L. & Wigfield, A. (1999). Dimensions of children's motivation for reading and their relations to reading activity and reading achievement. *Reading Research Quarterly, 34,* 452-477.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives*. *Handbook 1. Cognitive domain*. New York: David McKay.

Brooks, G., Pugh, A.K., & Schagen, I. (1998). *Reading performance at nine.* Berks.: NFER.

Brooks, G., Schagen, I., & Nastat, I. (1998). *Trends in reading at eight: A report on the 1995 survey of reading attainment in Year 3 in England and Wales.* Berks.: NFER.

Campbell, J.R., Kelly, D.L., Mullis, I.V.S., Martin, M.O., & Sainsbury, M. (2001). *Framework and specifications for PIRLS assessment 2001—2nd Edition.* Chestnut Hill, MA: Boston College.

Cochran, W.G. (1977). *Sampling techniques (3rd ed.).* New York: Wiley.

Cosgrove, J. (2003). *PISA 2000 unpublished student questionnaire data analyses for Ireland.* Dublin: Educational Research Centre.

Cosgrove, J., Kellaghan, T., Forde, P., & Morgan, M. (2000). *The 1998 assessment of English reading (with comparative data from the 1993 assessment).* Dublin: Educational Research Centre.

Cosgrove, J., & Morgan, M. (2000). The leisure activities of fifth grade pupils and their relationships to pupils' reading achievements. *Irish Journal of Education, 31,* 50-62.

Dickinson, D.K., & Tabor, P.O. (2001). *Beginning literacy with language: Young children learning at home and school.* Baltimore, Maryland: Paul H. Brookes.

Educational Research Centre (2002). *Drumcondra Sentence Reading Tests, Level 1, Forms A and B.* Dublin: Author.

Elley, W.B. (1992). *How in the world do students read? IEA Study of Reading Literacy.* The Hague: International Association for the Evaluation of Educational Achievement.

Elley, W.B. (1994). *The IEA Study of Reading Literacy: Achievement and instruction in thirty-two school systems.* Oxford: Pergamon.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behaviour: An Introduction to theory and research.* Reading, M.A: Addison Wesley.

Gambrell, L.B., Palmer, B.M., Codling, R.M., & Mazzoni, S.A. (1996) Assessing motivation to read. *The Reading Teacher, 49,* 518-533.

Ganzeboom, H.B., & Treiman, D.J. (1996). Internationally comparable measures of occupational status for the 1988 international standard classification of occupations. *Social Science Research*, *25,* 201-239.

Guthrie, J.T. & Wigfield, A., (2000). Engagement and motivation in reading. In M.L. Kamil, P.B. Mosenthal, P.D. Pearson, & R. Barr (eds.), *Handbook of Reading Research (Vol. 3)* (pp. 403-422). Mahwah, NJ: Lawrence Erlbaum Associates.

Harris, T.L., & Hodges, R.E. (1995). *The literacy dictionary: The vocabulary of reading and writing.* Newark, DE: International Reading Association.

Hidi, S., & Harackiewicz, J.M. (2000). Motivating the academically unmotivated: A critical issue for the 21st Century. *Review of Educational Research, 70,* 151-179.

Hoffman, J.V. (1986). *Effective teaching of reading: Research and practice*. Newark, DE: International Reading Association.

Kellaghan, T., Sloane, K., Alvarez, B., & Bloom, B. S. (1993). *The home environment and school learning*. San Francisco: Jossey-Bass.

Kirsch, I. (2001). The international adult literacy survey (IALS): Understanding what was measured. *ETS Research Report RR-01-25.* Princeton, NJ: ETS.

Kirsch, I., deJong, J., Lafontaine, D., McQueen, J., Mendelovits, J., & Monseur, C. (2002). *Reading for change: Performance and engagement across countries. Results from PISA 2000.* Paris: OECD.

Kish, L. (1957). *Survey sampling.* New York: Wiley.

Kispal, A., Gorman, T.P., & Whetton, C. (1989). *Reading Ability Series Levels A-F, teachers' handbook.* Windsor: NFER-NELSON.

Lumsden, L. (1997). *Expectations for students.* ERIC Digest Oregon: ERIC Clearinghouse on Educational Management.

McKenna, M. C., Kear, D. J., & Ellsworth, R. A. (1995). Children's attitude toward reading: A national survey. *Reading Research Quarterly, 30,* 934-956.

Martin, M.O., & Kelly, D.L. (1996). *TIMSS technical report volume I: Design and development.* Chestnut Hill, MA: Boston College.

Martin, M.O., & Morgan, M. (1994). Reading literacy in Irish schools: A comparative analysis. *Irish Journal of Education, 28.*

Martin, M.O., Forde, P.D., & Hickey, B.L. (1991). *Assessment of reading comprehension in primary school. A development project. Summary report.* Dublin: Educational Research Centre.

Mislevy, R. J., & Bock, R. D. (1990). *BILOG*. Mooresville, IN: Scientific Software.

Mullis, I.V.S., Martin, M.O., Gonzalez, E.J., & Kennedy, A.M. (2003). *PIRLS 2001 international report: IEA's study of reading literacy achievement in primary schools.* Chestnut Hill, MA: Boston College.

Mullis, I.V.S., & Smith, T.A. (1996). Quality control steps for free-response scoring. In M.O. Martin and I.V.S. Mullis (Eds.), *Third International Mathematics and Science Study: Quality Assurance in Data Collection*, pp. 5-1-5-29. Chestnut Hill, MA: Boston College.

Murray, T.S., Kirsch, I.S., & Jenkins, L.B. (1998). *Adult literacy in OECD countries: Technical report on the first International Adult Literacy Survey.* Washington, DC: US Department of Education, National Center for Education Statistics.

National Academy of Education (1993). *Setting performance standards for student achievement: A report of the National Academy of Education Panel on the evaluation of the NAEP trial state assessment: An evaluation of 1992 achievement levels.* Stanford, CA: National Academy of Education.

National Council for Curriculum and Assessment (NCCA) (1999). *The primary school curriculum: English.* Dublin: Author.

National Economic and Social Forum (NESF) (2001). *Lone parents.* Dublin: Author.

National Reading Panel (2000). *Report on the national reading panel – teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction.* Washington DC: National Institute of Child Health and Human Development.

OECD (Organisation for Economic Cooperation and Development) (1999). *Measuring student knowledge and skills: A new framework for assessment.* Paris: Author.

OECD (Organisation for Economic Cooperation and Development) (2001). *Knowledge and skills for life: First results of PISA 2000.* Paris: Author.

Pintrich, P.R., & deGroot, E. V. (1990). Motivational and self-regulated components of classroom academic performance. *Journal of Educational Psychology, 82,* 33-40.

Postlethwaite, T.N., & Ross, K.N. (1992). *Effective schools in reading: Implications for educational planners. An exploratory study.* The Hague: International Association for the Evaluation of Educational Achievements.

Raudenbush, S., Bryk, A., Cheong, Y.F., & Congdon, R. (2000). *HLM 5: Hierarchical linear and nonlinear modelling.* Lincolnwood, IL: Scientific Software International, Inc.

Sammons, P., Thomas, S., & Mortimore, P. (1997). *Forging links: Effective schools and effective departments.* London: Paul Chapman.

Schiefele, U. (1999). Interest and Learning from text. *Scientific Studies of Reading, 3,* 257-280.

Schunk, D. H., & Zimmerman, B. J. (1994). *Self-regulation of learning and performance: Issues and educational applications.* Hillsdale, N. J.: Erlbaum.

Schunk, D. H., & Zimmerman, B. J. (1997). Developing self-efficacious readers and writers: The role of social and self-regulatory processes. In J.T. Guthrie, & A. Wigfield (eds.), *Reading Engagement: Motivating readers through integrated instruction.* (pp. 34-50). Newark, D.E: International Reading Association.

Schmitt, M. C. (1990). A questionnaire to measure children's awareness of strategic reading processes*. The Reading Teacher, 44,* 454-461.

Sharp, C., Keys, W., & Benefield, P. (2001). *Homework: A review of recent research.* Sough, Berks.: National Foundation for Educational Research.

Shiel, G., Cosgrove, J., Sofroniou, N., & Kelly, A. (2001). *Ready for life? The literacy achievements of Irish 15-year olds with comparative international data.* Dublin: Educational Research Centre.

Shiel, G., & Kelly, D. (2001). *The 1999 National assessment of mathematics achievement.* Dublin: Educational Research Centre.

Shiel, G., & Murphy, R. (2000). *Drumcondra English Profiles: A framework for assessing oral language, reading, and writing in primary schools.* Dublin: Educational Research Centre.

Taylor, B.M. (2002). *Teaching reading: Effective schools, accomplished teachers.* Mahwah, NJ: Lawrence Erlbaum Associates.

Wigfield, A. (1997). Children's motivations for reading and reading engagement. In J.T. Guthrie, & A. Wigfield (eds.), *Reading engagement: Motivating readers through integrated instruction.* (pp.14-33). Newark, DE: International Reading Association.

Wixson, K. K., & Peters, C. W. (1984). Reading redefined: A Michigan Reading Association position paper. *Michigan Reading Journal*, *17*, 4-7.

Wray, D., Medwell, J., Poulson, L., & Fox, R. (2001). *Teaching literacy effectively in the primary school.* London: Routledge-Falmer.

Zimowski, M.F., Muraki, E., Mislevy, R.J., & Bock, R.D. (1996). *BILOG MG.* Chicago, IL: Scientific Software International.

APPENDIX A
NAER 2004 NATIONAL COMMITTEE MEMBERS

Name		Organisation
Éamonn	Murtagh	Department of Education and Science, Chair
Bairbre	Boylan	National Council for Curriculum and Assessment
Richard	Byrne	Department of Education and Science
Jude	Cosgrove	Educational Research Centre
Patrick	Forde	Educational Research Centre
Mary	Killeen	National Parents' Council-Primary
Ruby	Morrow	Church of Ireland Board of Education
Carmel	NicAirt	An Fóras Patrúntachta
Deirbhile	NicCraith	Irish National Teachers' Organisation
Niamh	Ní Mhaoilín	Gaelscoileanna
Helena	O'Connell	Catholic Primary Schools Management Association
Gerry	Shiel	Educational Research Centre
Áine	Sotscheck	Educate Together
Hannah	Wardle	Educational Research Centre

APPENDIX B
ANALYSIS OF CURRENT ENGLISH TEXTBOOKS

Introduction

Although the Primary School Curriculum makes reference to a range of text types that pupils should be engaging with as they progress through school, there is no explicit mention as to the relative emphasis that different text types are to receive. Thus, a textbook analysis was conducted with a view to gathering information on the proportion of text devoted to each of the two purposes (i.e., literary experience vs. acquiring and using information) and the proportion of continuous/non-continuous texts in the current English textbook series. Information on the presentation of the texts (e.g., length of passages, number and function of pictures accompanying passages) was also sought, since such information would be of assistance in developing passages which are authentic and representative of the material read by first/fifth class pupils, at least in their English classes.

At first class level, three series of English textbooks were analysed. All series included more than one textbook, and in most cases, each of the textbooks comprised several units (i.e., different short stories or chapters in literary books and different topics in informational books).

For each series, texts were analysed using a unit-level approach. A unit is defined as a piece or pieces of text of any kind which form a unified chunk of information. For each unit, text was classified by broad text type (e.g., continuous, non-continuous or mixture of continuous text accompanied by non-continuous text). The number of words devoted to literary text and informational text was also calculated at unit-level and book-level, and then aggregated across all the books in a series to get an estimate of the relative emphasis placed on each purpose within a series. Finally, the number of words devoted to each of the purposes was aggregated across the three series to ascertain the relative emphasis placed on literary and informational text in all first class English textbooks.

Once the purpose of each text was ascertained, texts were further classified (again using a unit-level approach) to determine the specific text type. All literary texts were classified as narrative. However, informational texts were further classified according to whether they were narrative, descriptive, expository, injunctive/instructional, argumentative, or a hybrid of these types, and word counts were used to determine the relative emphasis placed on each of these text types. Non-continuous informational text was also classified by text type (e.g., reference materials, forms, advertisements, charts and graphs, diagrams/schematics, tables and matrices, maps). However, the nature and dearth of non-continuous text made word counts difficult, and instead the instance of non-continuous text was noted (e.g., number of maps, diagrams, etc).

To get an idea of the presentation of texts, the actual number of words per page was calculated. The word count for each unit was calculated to get an idea of the average length of each passage. Finally, unit word counts were aggregated to ascertain the average number of words per book (for example, to ascertain whether books in a series were of increasing difficulty). The number and function of pictures (e.g., to aid comprehension or to break up the text) in each unit was also noted. To guide the development of the style and layout of TARA 04, a note of illustrations, font type, font size and spacing was made.

Finally, the general themes and topics addressed in the literary and informational texts were noted, and popular authors recorded.

Results of Analysis of Current First Class English Textbooks

Series 1

The series includes three books for literary experience, and two informational books. A word count indicates that 65% of text is for the purpose of literary experience, while 35% is devoted to informational text.

Literary experience: 91 pages, 4990 words.
Informational: 90 pages, 2800 words.

Characteristics of books for literary experience

Two of the literary books (33 and 45 pages) are comprised of short stories, and the remaining book is 18 pages long and contains one continuous story broken down into 7 chapters. There are between 160-300 words per unit (i.e., short story), although there is a wide range, with some containing as few as 50 words, others up to 400. For the first two texts, the amount of text per page was approximately 40 to 55 words (with a range of 25-65 per page). The exception was the final text, in which there are approximately 80 words per page. All pages of text have at least one illustration per page (or one picture per 50 words) which appears to serve the function of aiding meaning (e.g., illustrating information in the text).

Characteristics of informational books

The two books, both 45 pages, are also broken down into units, each addressing a different topic. There are approximately 140 words per unit in the first book and 210 words per unit in the second book (although there was a wide range, between 85-330). The word count per page ranged widely, with approximately 25-40 words per page in the two different books.

The type of text was mixed. Units combined narrative text (e.g., facts presented in a story format) with expository and descriptive text. It was common in each unit to combine descriptive text with both pictures which were more decorative and labelled with pictures to aid comprehension. The labelled pictures varied from a simple picture with a label of what the picture was, to a picture labelling different parts of the object pictured.

Topics covered included animals and nature, transport, and personal health and safety.

Series 2

The series comprised three books for literary experience, and two informational books. A word count indicates that 75% of text is for the purpose of literary experience, while 25% is devoted to informational text.

Literary experience: 78 pages, 3900 words.
Informational: 28 pages, 1425 words.

Characteristics of books for literary experience

All three of the literary books are between 21 and 29 pages and consist of two or three short stories. There are 330 words per unit (i.e., per short story) in the first book, 400 words per unit in the second book and 850 words per unit in the third book, although there is a wide range per unit, with some containing as few as 270 words, others up to 1000. The amount of text per page ranges from 40 to 60 words. All pages of text have at least one illustration per page (or one picture per 50 words). These appear to serve the function of aiding meaning (e.g., illustrating information in the text).

Characteristics of informational books
The two books contain 15 pages each. There are approximately 630 words in the first book, 800 in the second. The word count per page was approximately 50 words per page.

The type of text was mixed. The books combined expository and descriptive text. It was common in each unit to combine descriptive text with both pictures which were more decorative and labelled pictures to aid comprehension. The labelled pictures were generally simple, primarily stating what the picture was (although one picture labelled parts).

Unlike in Series 1, each of the informational books in Series 2 tended to cover two themes in-depth, with different aspects of sleep and communication covered.

Series 3
The series comprises three books written for the purpose of literary experience. The first and second books, each approximately 50 pages, consist of 6 to 7 short stories. The first book has approximately 340 words per unit, and 50 words per page, and the second has approximately 580 words per unit, 80 words per page (although the word count ranges from 40-100 per page and 290-690 per unit across the two books). The third book consists of one story, and is 28 pages, with approximately 35 words per page. In the three books, all pages of text have at least one illustration per page which appears to serve the function of aiding meaning (e.g., illustrating information in the text).

Literary experience:122 pages, 7070 words

Series 3 does not include any informational books in the core reading series, although supplementary material and sets of parallel books are available, which include both fiction and non-fiction books. These were not included in the analysis.

Summary of results: first class
Based on word counts in each book, when all three series are included, 80% of the text is narrative short stories for literary purposes, and 20% is informational text. Omitting the series without the information text, averages of 70% and 30% are obtained. Common themes found in the first class texts included animals or monsters, fantasy, books and reading, playing, sleeping and transport. The characteristics of the literary books at first class level are summarised in Table A1.

Table A1 Summary of characteristics of books written for literary experience, across the three series of first class texts

		Words per unit		Words per page	
		Mean	*Range*	*Mean*	*Range*
Series 1	Book 1	160	50-290	40	25-45
	Book 2	300	225-425	55	45-65
	Book 3	300*	-	80	25-135
Series 2	Book 1	330	310-360	50	45-50
	Book 2	400	275-530	40	25-55
	Book 3	850	700-1060	60	25-110
Series 3	Book 1	340	290-460	50	40-65
	Book 2	580	430-835	80	65-100
	Book 3	920**	-	35	15-60
Average over 3 series		460	50-1060	55	35-75

*This story is one continuous story of 1500 words, but there are approximately 300 words per *chapter*.
**This book consists of one continuous story.

Each of the three series included three texts of increasing difficulty (i.e., one with fewer pages/words per page, and the next slightly more pages/words per page, with the third book in two of the three series consisting of one continuous story, and in one series of two longer stories). Generally there were pictures on each page (e.g., a picture for every 50 words) designed to assist the reader in constructing the meaning of the text. The characteristics of the informational books at first class level are summarised in Table A2.

Table A2 Summary of characteristics of information books, across the three series of first class texts

		Text type		Words per unit		Words per page	
		Narrative	*Expository / descriptive*	*Mean*	*Range*	*Mean*	*Range*
Series 1	Book 1	Hybrid (narrative/exp/desc)		140	90-190	25	18-35
	Book 2	Hybrid (narrative/exp/desc)		210	140-330	40	25-45
Series 2	Book 1	None	100%	40	10-130	45	15-80
	Book 2	None	100%	80	40-90	55	40-90
Average over 3 series		**-**	**-**	**115**	**40-330**	**40**	**18-90**

*One series did not include any informational texts in its core reading.

Informational texts at first class level contain mostly continuous text but are accompanied by labelled pictures to illustrate concepts discussed in the text. Non-continuous text was virtually absent in the three series (although the website for Series 3 notes that their skills books provide opportunities for functional reading, e.g., timetables and recipes). The closest to non-continuous text were simple pictures with labels, although this could not really be considered a 'diagram', since the definition for diagram is a graphic that *explains* rather than represents, and these pictures just represent. However, each of the readers did contain a Table of Contents, and the two informational books in Series 1 also contained a picture index (which is consistent with the Primary School English Curriculum where it is suggested children at this level should become familiar with Table of Contents and Indices).

It should be noted that it was sometimes difficult to distinguish between expository and descriptive text. The Literacy Dictionary (Harris & Hodges, 1995) notes that expository text may overlap with the other forms of informational text (description, argumentation, narration). The dictionary notes that descriptive text gives a picture description of a character (physical object) or event, and expository is intended to explain or set forth information, and may use description in doing so. For example, in

the 'Sleep' book (Series 2), the majority of the text is expository in describing different aspects of sleep, such as how long and why we sleep, hours of sleep, where we sleep, etc., but in conveying information about sleep, some description is used, such as descriptions of bed types. Similarly, in Series 2's 'Hello' book the majority of text describing communication is expository in style but there is some physical description of phone cards and telephones. This difficulty was resolved by recording these text types as expository/descriptive hybrids.

Results of Analysis of Current Fifth Class English Textbooks

At fifth class, four series[8] of English texts were analysed. All but one series included more than one textbook, and each of the textbooks comprised several units (i.e., different short stories or chapters in literary books and different topics in informational books). The procedures used for the fifth class analysis are the same as those used for the first class analysis.

Series 1

The series includes one book for literary experience, and one informational book. A word count indicates that approximately 65% of text is for the purpose of literary experience, while 35% is devoted to informational text.

Literary experience:149 pages, 30600 words.
Informational: 96 pages, 16000 words.

Characteristics of book for literary experience

The literary book consisted of 20 short stories. There are approximately 1500 words per unit (i.e., per short story), although there is a wide range per unit, with some containing as few as 1000 words, others up to 3000. The amount of text per page was approximately 200 words, although there was a range (from 170 to 240 words). The majority of pages of text have at least one illustration per page, although the pictures tended to be decorative rather than aiding comprehension.

The book contains extracts from the following authors: Jean Ure, Morgan Llewelyn, CS Lewis, John Quinn, Mary Norton, Ann Carroll, Betsy Byars, Eilis Dillon, Berlie Doherty, Maria Conlon-McKenna, Oscar Wilde, Gregory Maguire, Bridget Dirrane, William Collins, Anna Sewell and Maeve Friel.

Characteristics of informational book

The informational book consisted of 12 units, each addressing different topics. Topics were varied and included nature, history and geography, people and culture, art, personal health and safety, and transport.

Each unit had 8 pages, and there were on average 1350 words per unit, although this varied from 735 to 1860 words. There were approximately 170 words per page, although this varied from 90-230 words per page.

The type of texts used were mixed. Approximately 25% of informational text was narrative (i.e., where facts were presented in a story format), and 75% was descriptive-expository. As was the case with first class informational texts, it was difficult to distinguish between expository and descriptive text. PISA texts also reflect the overlap between expository/descriptive as some passages in PISA 2000 were classified as 'descriptive-explanatory'.

The majority of text consists of descriptive-expository text accompanied by decorative pictures (i.e., where they served to break up text and make the page more

[8] One publishing company produces English series for 3rd to 6th class only so there is one more series for fifth class than for first class.

interesting), photographs and simple labelled pictures (to aid comprehension), three diagrams (e.g., pictures where the function is to explain a process rather than simply represent), and eight maps.

Series 2

The series includes two books for literary experience, and one informational book. Word counts indicate that approximately 55% of text is for the purpose of literary experience, while 45% is devoted to informational text.

Literary experience:127 pages, 21500 words.
Informational: 74 pages, 17600 words.

Characteristics of books for literary experience
The literary books were approximately 65 pages long, and were split into 8 to 10 different units (i.e., short stories). The length of the stories varied from 3 to 11 pages. There were approximately 1200 words per unit (although this varied from 550-2000 per unit), and 170 words per page (although this varied from 130-200 words per page). The stories tended to have one picture for every two pages, and the function of these pictures was decorative.

The book contains extracts from the following authors: Norton Juster, Jacqueline Wilson, Judith Kerr, JRR Tolkien, Michael Morpurgo, Roald Dahl, Chris d'Lacey, EB White, Isaac Asimov, Yvonne MacGrory, Walter Buchignani, Michael Scott, Phillippa Pearce, Creina Mansfield, Antoine de Saint-Exupery, E Nesbit and Betsy Byars.

Characteristics of informational book
Rather than focusing on different topics, the informational book was broken down into sources of information (e.g., letters, diary entries, instructional manuals, etc). Within each source, a range of topics was examined, including lives of influential people, arts and entertainment, sports and hobbies, science, history and geography, and personal health and safety.

There were on average 600 words per unit, although this varied from 275 to 1750 words. There were approximately 235 words per page, although again this varied from 140-325 words per page.

A range of text types, both continuous and non-continuous, was used. The types of continuous texts used were narrative, expository, descriptive, and injunctive, and in some cases a hybrid of texts. The emphasis tended to be placed on narrative and a mix of expository/descriptive/ injunctive texts.

- Narrative text: ~7600 words – 46%
- Instructive/injunctive text: ~1350 words – 8%
- Expository/descriptive/injunctive hybrid: ~4810 words – 46%

The types of narrative text used included biography and autobiography, letters, diary entries, speeches, news reports, and interviews.

Injunctive text included general instructions presented in continuous sentences/paragraphs, instructions for making arts and crafts, and rules and regulations (e.g., list of do's and don'ts).

Narrative texts generally included decorative pictures and photographs. The other text types were often combined with non-continuous texts. Injunctive texts were accompanied by checklists, pictures to aid comprehension (visual instructions), and decorative pictures. Expository and hybrid texts were accompanied by decorative

pictures, diagrams, maps, and/or checklists. The book also contains a Table of Contents and glossary.

Series 3

There is only one book in this series. The book consists of six stories written for literary experience (e.g., fiction stories where the purpose is to entertain and engage the imagination). These stories range from 7 to 13 pages, with approximately 2400 words per story (with a range of 1400-2800), and 230 words per page (with a range of 200-245).

The book also includes 5 stories which are based on true stories and facts, and relay facts and historical events in a narrative format. These stories ranged from 9 to 15 pages, with approximately 2300 words per story, and 210 words per page. There are also 5 two-page comics which presented information in a narrative format.

Based on the above categorisation:
Literary experience: 64 pages, 15000 words: 54%.
Informational (including comics): 63 pages, 13000 words: 46%.

The topics addressed in the informational stories were in the areas of history (ancient people, historical events), science (fossils and dinosaurs), and personal health and development. All informational text presented in this book was in narrative format. There were no examples of the other types of informational text (e.g., expository, descriptive), nor were there any non-continuous texts (with the exception of the table of contents). There was approximately one picture for every two pages, and these appeared to be more decorative in their function.

The book contains a number of purpose written texts (since no authors are noted).

Series 4

The series includes one book for literary experience, and one informational book. Word counts indicate that approximately 75% of text is for the purpose of literary experience, while 25% is devoted to informational text.

Literary experience: 105 pages, 28000 words.
Informational: 52 pages, 10700 words.

Characteristics of book for literary experience
The literary book comprised 18 units (i.e., stories). The length of the units varied from 4 to 11 pages. There were approximately 1600 words per unit (although this varied from 750-3000 per unit), and 270 words per page (although this varied from 185-400 words per page). The stories tended to have one picture for every few pages, and the function of these pictures was more decorative than instructive (e.g., to break up the text).

The book contains excerpts from the following authors: Gene Kemp, Morris Gleitzman, Lois Lowry, Paul Jennings, EB White, Willis Hall, Michael Morpurgo, Gillian Cross, Vivien Alcock, JK Rowling, Eilis Dillon, Philip Pullman, John Quinn, Malorie Blackman, Betsy Byars, Elizabeth O'Hara, Katherine Paterson, Oscar Wilde.

Characteristics of informational book
The informational book consisted of 19 units, each addressing different topics.
Topics covered include: history, personal health, science and nature, geography, inventions, sports, art and entertainment, and lives of influential people.

There were on average 540 words per unit, although this varied from 220 to 1150 words. There were approximately 200 words per page, although again this varied from 110-300 words per page.

A range of text types, both continuous and non-continuous, was used. The types of continuous texts used were narrative, expository, descriptive, and injunctive, and in some cases a hybrid of texts. For the most part, the text was expository, but in elaborating on concepts, there was some description involved). Overall, the emphasis tended to be placed on expository text:

- Narrative text: ~1800 words, 18%
- Expository/descriptive text: ~8900 words, 80%
- Instruction/injunctive: ~170 words, 2%

Narrative text was accompanied by photographs and included biography, news reports, and interviews. Expository/descriptive text was often accompanied by decorative pictures, labelled pictures, photographs, and in many cases was combined with non-continuous text (tables, diagrams, maps or checklists). The book also contained examples of non-continuous text on their own (i.e., not combined with expository text – a timeline, table of contents, glossary and index). The book itself also had its own table of contents, index and glossary.

Summary of results: fifth class

Table A3 shows a summary of characteristics of fifth class English texts across the four series based on word counts in each book. Table A4 shows some additional information about the literary texts, while Table A5 shows some additional information about the informational texts.

Table A3 Summary of breakdown of the proportions of texts written for literary experience and information purposes, across the four series of fifth class texts

	Purpose	
	% Literary Experience	*% Informational*
Series 1	65%	35%
Series 2	55%	45%
Series 3	55%	45%
Series 4	75%	25%
Average over 4 series	**65%**	**35%**

Table A4 Summary of characteristics of books written for literary experience, across the four series of fifth class texts

	Words per story/unit		**Words per page**	
	Mean	*Range*	*Mean*	*Range*
Series 1	1500	1000-3000	200	170-240
Series 2	1200	550-2000	170	130-200
Series 3	2400	1400-2800	230	200-245
Series 4	1600	750-3000	270	185-400
Average over 4 series	**1675**	**550-3000**	**220**	**130-400**

The most popular writer was Betsy Byars whose work was included in 3 series (7 stories in total). The following authors were also popular (i.e., their work featured in at least 2 of the 4 series): EB White, Michael Morpurgo, Eilis Dillon, John Quinn, and Oscar Wilde.

Table A5 Summary of characteristics of books written for literary experience, across the three series* of fifth class texts

	Text type		**Words per unit**		**Words per page**	
	Narrative	*Expository / descriptive / injunctive*	*Mean*	*Range*	*Mean*	*Range*
Series 1	25%	75%	1350	735-1850	170	90-230
Series 2	47%	53%	600	275-1750	235	140-325
Series 4	20%	80%	540	220-1150	200	110-300
Average over 3 series	**30%**	**70%**	**830**	**400—1600**	**200**	**110-285**

*Series 3 consisted of one book for both informational/literary purposes and all informational text was narrative. There were approximately 2400 words per unit, 230 words per page. Since the structure of this series was different from the three other series it was not included in the table.

Although the TARA booklets have items on reading timetables/programmes, bar charts, dictionaries, and weather reports, there was an absence of non-continuous texts such as these in the fifth class textbooks. Although not reflected in textbooks, such skills are emphasised in the Primary School English Curriculum (NCCA, 1999).

APPENDIX C
SAMPLE READING TASKS FROM NAER 2004 PILOT SURVEY

Introduction

The passage extracts and questions here are taken from the NAER pilot study. These passages were piloted but not selected for the final versions of the test. Examples provided represent each type of text contained within the study; narrative, expository and documents (or literary experience, informational/continuous, and informational/non-continuous), and also demonstrate how each relevant process is tapped using retrieve, inference, and interpret example questions. Percent correct is provided for each example question. However, note that percent correct is unweighted, and based on the pilot survey sample, and not necessarily representative of the population of all first and fifth class pupils. Two of the three fifth class passages have been shortened to excerpts since the entire text is quite long. Fifth class pupils were also asked some evaluative questions, though none are available from the pool of dropped items. For examples of evaluative reading processes please refer to Figure 3 on page 8.

First Class Sample Tasks

All first class texts and questions were presented in a format and font size similar to the main reading schemes used at this class level. Text was also accompanied by pictures wherever possible.

Mill Street Party: *Literary Experience/Continuous Text*

Every summer the people of Mill Street have a big street party.
Everyone is invited.
The party starts when the band marches up and down the street.
'BOOM BOOM' goes the drum!
Everyone claps their hands to the music.
After that it's time to eat.
Plates of sausages, chips and pizza are passed around.
The children have their faces painted.
Later on the magician arrives.
He does lots of magic tricks.
Everybody wants to see him pull a rabbit out of a hat.
The party is over by bedtime.

Note: An asterisk (*) denotes the correct answer.

Retrieval question:	The big street party is held…	***% endorsing***
	a. in the spring	6
	b. in the summer	83*
	c. in the autumn	3
	d. in the winter	2
Inference question:	A magician is someone who…	
	a. makes pizza	6
	b. paints pictures	4
	c. plays the drum	12
	d. does tricks	69*
Interpret question:	What is the best name for this story?	
	a. The Magician	10
	b. The Band Plays Again	8
	c. The Street Party	61*
	d. Summer is Here	10

New Puppy: ***Informational/Continuous Text***

Sara got a new puppy!
The man at the pet shop gave Sara this note about how to keep her puppy safe and healthy.

HOW TO CARE FOR YOUR NEW PUPPY!

Puppies need a doctor, just like you do.
Take your new puppy to the vet to make sure that he is healthy.

You should give your puppy a bath every week.
This will help keep your puppy's coat clean and shiny.

Get a collar and lead for your puppy so that you can take him for walks.

Note: An asterisk (*) denotes the correct answer.

Retrieval question:	When you take your puppy for walks you should use…	***% endorsing***
	a. a lead.	77*
	b. a vet.	5
	c. toys.	4
	d. a ball.	12
Inference question:	To keep your puppy's fur clean and shiny you should…	
	a. walk your puppy every week.	6
	b. wash your puppy every week.	80*
	c. put a name tag on your puppy.	6
	d. take your puppy to the vet.	5
Interpret question:	How do you think Sara feels about getting a new puppy?	
	a. worried	7
	b. sad	11
	c. scared	4
	d. excited	75*

Food Table of Contents: ***Informational/Non-continuous text:***

Making Food

Contents

Chapter 1	Making a sandwich
Chapter 2	Baking a pie
Chapter 3	Cooking vegetables
Chapter 4	Cooking meat
Chapter 5	Frying fish

Note: An asterisk (*) denotes the correct answer.

Retrieval question:	In Chapter 1 you would find out how to…	***% endorsing***
	a. ...make a sandwich.	79*
	b. …cook vegetables.	6
	c. …cook meat.	7
	d. …fry fish.	4

Inference question:	Which chapter would help you learn about cooking chicken?	
	a. Chapter 1	6
	b. Chapter 2	12
	c. Chapter 4	70*
	d. Chapter 5	8

Interpret question:	Where would you use this book?	
	a. in the kitchen	70*
	b. in the garden	8
	c. in the garage	3
	d. in the bedroom	16

Fifth Class Sample Tasks

The fifth class passages appeared in varied format and font to provide more variety throughout the test, as at this level these aspects are not expected to influence reading comprehension. In the case of the first two passages, extracts rather than the entire text is shown here. The informational non-continuous text example shown here is, unfortunately, not very representative of the other non-continuous materials in the fifth class test, which includes timetables, a map, data tables, TV guides, etc. However it is the only available sample text at this level.

Excerpt from I Am David: *Literary Experience/Continuous Text*

[Total word count in test booklet = 1013]

David is hiding from the concentration camp guards. He has a dog as a friend and companion and is hoping his newfound freedom might last as he attempts to find safety. But his hopes are beginning to fade...

…when they called to him, he would run, and then the shot would be fired which had been waiting for him ever since that night when he had walked calmly towards the tree on the way to the mine outside the camp. But this time he would not be able to walk calmly from them. He now knew how wonderful life could be, and his desire to live would spur him on. He would run – he knew it. And it would be a victory for *them*.

David remembered all the pain and bitterness he had ever known – and how much he could remember in such a short time! He recalled, too, all the good things he had learned about since he had gained his freedom – beauty and laughter, music and kind people, Maria, and a tree smothered in pink blossom, a dog called King to walk by his side, and a place to aim for…

This would be the end. He pressed his face into the dog's long coat so that no one should hear him, and wept. He wept quite quietly, but the dog grew uneasy and wanted to whimper again.

David stopped crying. 'God,' he whispered, 'God of the green pastures and still waters, I've one promise of help left, but it's too late now. You can't do anything about this. I don't mean to be rude, because I know you're very strong and you could make those men down there want to walk away for a bit. But they won't. They don't know you, you see, and they're not afraid of you. But they are afraid of the commandant because he'll have them shot if they leave their posts. So you can see there's nothing you can do now. But please don't think I'm blaming you. It was my own fault for not seeing the danger in time. I shall run… Perhaps you'll see they aim straight so it doesn't hurt before I die. I'm so frightened of things that hurt. No, I forgot. I've only one promise of help left, and it's important you should help the dog get away and find some good people to live with. Perhaps *they'll* shoot straight anyway, but if they don't it can't be helped: you must save the dog because it once tried to protect me. Thank you for having been my God: I'm glad I chose you. And now I must run, for if I leave it any longer I won't have the courage to die. I am David. Amen.'

Note: An asterisk (*) denotes the correct answer.

		% endorsing/correct
Retrieve question (constructed response):	What was the name of David's dog? *Answer: King.* Example incorrect responses were: "companion" and "sheepdog". These received no credit.	77
Inference question:	David discovered "how wonderful life could be"…	
	a. …when he was praying to God.	17
	b. …before he was a prisoner in the camp.	18
	c. …after he escaped from the camp.	53*
	d. …when he was a prisoner in the camp.	11
Interpret question:	According to David, the guards were most afraid of…	
	a. …their commandant.	64*
	b. …the dog.	20
	c. …God.	5
	d. …escaped prisoners.	10
Interpret question (constructed response):	Do you think that David felt he would be able to get away safely? a. Yes b. No. Give one example from the story to support your answer. *Answer: No must be circled or implied in response with plausible evidence from the text.* Example correct responses: 1) NO because he said in his pray that he would die. 2) NO because he thought he was going to be brought back to camp. Example incorrect responses: 1) NO – *Not sufficient for credit, evidence needed.* 2) YES because he escaped before and he could do it again. – *This is true, but it is not what* David *thought.* 3) YES because he had a dog by his side to help him. – *This is inaccurate and does not receive credit.*	58

Excerpt from Hanging Out With Chimps: ***Informational/Continuous Text***
[Total word count in test booklet = 685]

Sound familiar?

Chimpanzees are intelligent. They can be tender, and they can be brutal. Does that remind you of any other species? Yes, chimps and humans are alike in some amazing ways. For example, most chimpanzee mothers are protective, affectionate, and playful. So are older brothers and sisters. They help care for babies. Sometimes chimps even adopt orphans.

But chimps are not always caring. Mothers sometimes kill and eat others' babies. And in the 1970s, a deadly war broke out between chimp groups. A whole group got wiped out.

"When I first started at Gombe," Goodall said, "I thought the chimps were nicer than we are. But time has revealed that they are not." Like humans, chimpanzees are good, bad, and complicated.

Tomorrow at Gombe

Gombe Stream Research Centre began as one woman with guts and binoculars. Now it includes full-time researchers and visiting scientists. Before sunrise tomorrow, they'll head back into the forest for more chimp-watching.

Jane Goodall probably won't join them. Since 1985 she's spent most of her time on the road. She tells people around the world about chimpanzees—and the need to save them.

taken from The National Geographic Kids' Website:
http://www.nationalgeographic.com/kids/

		% endorsing
Retrieve question:	What happened in Gombe in the 1970s?	
	a. A deadly war broke out between chimp groups	94*
	b. Jane Goodall opened the Gombe Stream Research Centre	2
	c. A deadly was broke out between chimps and humans	2
	d. Jane Goodall went on the road to tell people about chimpanzees	2
Inference question:	According to the passage, Gombe researchers gather information on chimps by…	
	a. …watching chimps at the zoo in Tanzania.	15
	b. …watching chimps in their natural environment.	79*
	c. …asking the locals in Gombe about chimpanzees.	6
	d. …watching videos about chimpanzees.	1

		% correct
Interpret question (constructed response):	When Jane Goodall first started at Gombe she thought chimps were nicer than humans but then she realised chimps could be good bad and complicated. What do you think helped change her mind? *Answer: The deadly war between chimp groups and/or a reference elsewhere in the text to mother chimps killing and eating other chimps' babies.* Example correct responses: 1) I think the war between chimp groups helped change her mind. 2) Maybe she saw the mother eating one of the baby chimps. Example incorrect response: 1) Because humans are good bad and complicated and so are chimps.	57

Evolve Car Brochure: *Informational/Non-continuous Text*

Great new deals from **EVOLVE** car showrooms

Dublin — Belfast — Kilkenny — Galway — Cork — Drogheda — Tralee

Our new Spring range now includes:

- ***Evolve Caprice.*** This nippy little two-seater is available in white, red, blue, silver and black. Although compact, it has a large boot. And for €12 000, at 14 kilometres per litre and a 1.1 litre engine which uses unleaded petrol, you really can't go wrong. Power steering for easy handling. Five-year or 80,000 kilometre warranty. CD stereo system extra.
- ***Evolve Lida.*** This diesel-driven 1.8 litre beauty does 20 kilometres per litre and costs only €13 000. It will comfortably seat four and has plenty of boot space. Power steering not available.
- ***Evolve Space Mobile.*** The space mobile was built with your comfort in mind. For €15 000, it seats six. Its 2.4 litre engine does 15 kilometres to the litre and uses unleaded petrol. Power steering for even the most difficult driving conditions. Available in blue, silver, gold and black. Five-year or 70,000 kilometre warranty.
- ***Evolve Turbo X.*** This powerful little car is available in white, red, silver and black. With a 2.4 litre engine, power steering and sunroof, you have guaranteed driving pleasure. The Turbo X does an amazing 20 kilometres to the litre (unleaded petrol), and seats four people. The cost? Only €18 500. CD stereo system included.

		% endorsing/correct:
Retrieve question:	Which car does not have power steering? a. Caprice b. Space Mobile c. Lida d. Turbo X	 5 3 77* 7
Retrieve question:	How many colours does the Space Mobile come in? a. Three b. Four c. Five d. Six	 2 87* 4 1
Interpret question (constructed response):	In your opinion, which car would be most suitable for a large family? Give one reason for your answer. *Answer: The Space Mobile with reference to its large size.* Example correct responses: 1) Space Mobile because there's six seats. 2) Space Mobile, it's a people carrier - *allowed because size implied and shows knowledge of people carriers' capacity.* Example incorrect response: 1) Space Mobile. *Needs explanation to get credit.* 2) Turbo X because it's powerful.	81