Reading for the new ERA:

Assessing reading in a digital environment

Dara Searle, Tom Lumley, Juliette Mendelovits
Australian Council *for* Educational Research

ABSTRACT

In 2009 the Programme for International Student Assessment (PISA) includes an Electronic Reading Assessment (ERA) as an international option, in which 22 of the 67 main survey countries are participating. The inclusion of ERA in PISA 2009 reflects the fact that the reading of electronic (digital) texts is becoming increasingly necessary and prevalent: the skills and knowledge required for digital reading are essential for full, creative and reflective participation in the 21st century.

The ERA conceptualisation took as its starting point the assumption that it was to be an assessment of reading in a digital environment, rather than a computer-delivered assessment of reading in the print medium. Creating an assessment of electronic reading poses the challenge of providing an authentic measurement that represents what is essential to reading in any medium as well as what is different and unique about this medium. The paper will describe the construct of reading in the electronic medium that was developed for the ERA, in which text processing (common to print and electronic reading) and navigation (peculiar to electronic reading) are conceived of as crucial variables. The extent to which the interaction between these two variables explains item difficulty will be explored, using results from a predictive study of the difficulty of ERA field trial items conducted before data collection, together with the field trial results. The usefulness of this construct both in terms of guiding item development and in better understanding at a more general level the nature of reading in the digital medium will be discussed.

TEXT

The Programme for International Student Assessment (PISA) is a standardised assessment that is administered to 15-year-olds in schools. PISA surveys take place every three years. The first survey was implemented in 2000 in 43 countries, the second in 2003 in 41 countries, and the third in 2006 in 57 countries. In 2009, sixty-seven countries are participating in the fourth assessment, which is being conducted in Australia right now (July to September 2009). In all cycles, the domains of reading, mathematics and science are assessed. For each assessment, one of the three domains is chosen as the major domain and it is given greater emphasis, while the remaining two areas, the minor domains, are assessed less thoroughly. In 2000 the major domain was reading, in 2003 it was mathematics and in 2006 science. In 2009 reading is again the major domain, allowing a full review of the framework and the inclusion of new elements.

The definition of reading literacy for PISA 2009 is as follows:

> "Reading literacy is understanding, using, reflecting on and engaging with written texts, in order to achieve one's goals, to develop one's knowledge and potential, and to participate in society."

To represent the range of item difficulty, and to ensure broad coverage of the domain, the PISA framework defines several task characteristics that are used in the construction of all reading literacy tasks. These task characteristics are: situation (personal, public, occupational and educational), medium (print and electronic), environment (authored and message-based)[1], text format (continuous, non-continuous, mixed and multiple), text type (description, narration, exposition, argumentation, instruction and transaction[2]) and aspect (access and retrieve, integrate and interpret and reflect and evaluate). Within the aspect variable, *access and retrieve* items require students to locate information, *integrate and interpret* items require interpretation of a text and/or integration of information across a text or texts) and *reflect and evaluate* items draw primarily on outside knowledge, and ask readers to relate this to the text they are reading.[3]

[1] Note that the Environment classification only applies to texts in the electronic medium.
[2] Note that the text type "transaction" is only applied to texts in the electronic medium.
[3] Detailed definitions of each of these task characteristics are given in the PISA framework publications (OECD, 2000, 2006).

The first published version of the reading framework mentioned the possibility of including electronic texts in a "future cycle". That "future cycle" has arrived. The 2009 cycle of PISA includes an electronic reading assessment (ERA) as an international option. Twenty of the 67 main survey countries are participating in this option, and the incorporation of electronic reading is one of the major modifications to the new reading framework.

The rationale for the inclusion of an electronic reading assessment is clear. As of 2007, almost one and a half billion people – one-fifth of the world's population – were reading on line (International Telecommunications Union, 2009). As accessing information via networked computer technologies becomes the norm, skill in electronic reading becomes increasingly important in personal and social life. Electronic reading is necessary too in the occupational domain: recent research suggests that it is more often becoming essential for workers to use computers as part of their jobs (OECD and STATCAN, 2005).

While many of the skills needed to read an electronic text are similar to those needed to read a print text, differences between the print and electronic environment require readers to develop new skills.

The way in which electronic reading is defined in PISA recognises that electronic reading is not just reading print text on a computer screen. Three major differences between print and electronic texts are outlined below, each followed by a short description of how the new PISA reading framework and the ERA instrument have addressed them.

a) When compared with print reading, readers of electronic texts are more likely to engage with different kinds of texts from different sources.

The PISA ERA framework outlines the kinds of texts that are found in the digital medium and represents this diversity in the assessment instrument with mixed and multiple texts, where readers need to integrate information from several sites, or from pages presenting information in different forms.

b) There is a greater onus on the reader to evaluate the text. This is because electronic texts have not typically undergone the scrutiny that is usually involved in the publication of a print-based text.

The large amount of information available through the internet means that readers need to swiftly evaluate the credibility of information they see, increasing the demand on critical

thinking in the electronic medium. PISA ERA *reflect and evaluate* items have a strong focus on the probity, relevance and credibility of the stimulus material.

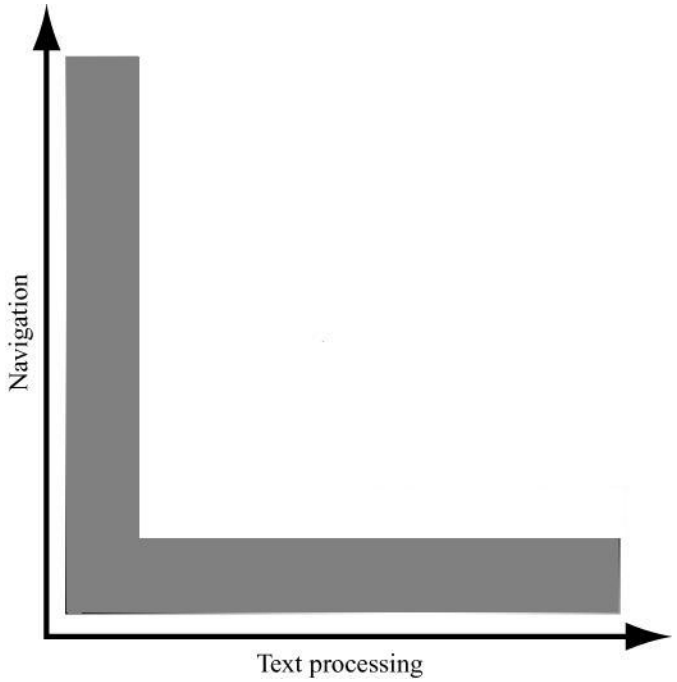
c) There is a greater onus on readers to select and construct the text. In print-based texts the physical status of the text encourages the reader to approach the content in a particular sequence. In contrast electronic texts have navigation tools and features that make it possible, and indeed even require, that readers create their own reading sequence.

The PISA framework has extended the definition of the *access and retrieve* aspect to acknowledge that the vast amount of information available in the electronic medium changes the nature of tasks requiring the retrieval of information. Readers more than ever need to be able to skim and search, and to navigate across oceans of information in a deliberate way. The ERA items provide examples of tasks that require construction of the reading text using both textual clues and navigation tools such as drop-down menus, embedded links and tabs.

Since readers of electronic texts typically need to work through multiple texts, of different kinds, which do not appear in a pre-defined sequence, perhaps the most salient difference between print and electronic reading is the unique way in which readers need to navigate their way through these texts, and to move with facility from one to another. Readers need to be able to scroll down a page, recognise and click on a link, and use back and forward buttons. The novelty of the electronic medium makes its navigation skills seem like something extrinsic to "reading" – "ICT literacy", perhaps. Yet these kinds of skills and knowledge are surely analogous to those needed in the print medium. When encountering print texts, readers need to know how to turn a page, how to use headings and other markers for orientation, and how to use a table of contents or an index. Because we are so used to the print medium, we are generally unaware that such navigation skills are part of the indispensable knowledge toolkit of print reading. The navigation knowledge required for the digital medium is just as much an essential part of electronic reading as knowing how to turn a page is indispensable to print reading. The PISA 2009 framework recognises that both navigation and text processing skills are typically required to complete electronic reading tasks. Some tasks place more emphasis on navigation. For example, a task with high navigational demands might require students to locate information by navigating across multiple web pages. Other tasks place more emphasis on text processing. For example, a task with high text processing demands might require students to interpret a long and complex

piece of text. The ERA instrument includes a pool of tasks that systematically vary the weighting of these two skills. The relationship between the two skills is represented in Figure 1 below.

Figure 1 Relationship between navigation and text processing in ERA tasks



In Figure 1 the horizontal axis represents the cognitive load that comes from processing the text, while the vertical axis represents the cognitive load that comes from the navigation required to successfully complete the task. In any single electronic reading task, one can represent the amount of navigation and text processing required and can imagine the task's position within Figure 1. Tasks that require low levels of both skills, that is, require little or no navigation and minimal text processing, would appear at the bottom left corner of the graph, close to the origin. Tasks that require high levels of both navigation and text processing would appear in the top right corner of the graph: the further from the origin, the more complex the task. In between these two extreme cases any combination of these two variables is possible. Tasks that require high levels of text processing, but low levels of navigation, would appear at the bottom right of the graph. A task of this kind might involve, for example, dealing with a text that is dense or complex (therefore requiring a high level of text processing) but that is immediately visible to the reader (thus requiring no navigation). It is considered necessary to include a small number of

tasks of this kind because although they do not require the skills that are unique to electronic reading, they do represent one kind of task that might be required in the real-life electronic environment. If this kind of task were excluded, the distance between print and electronic reading would be artificially inflated.

Conversely, the ERA assessment also includes tasks that require high levels of navigation, but low levels of text processing. These tasks would be represented in the top left quadrant of the graph in Figure 1. This kind of task might require the use of multiple strategies to navigate between web pages (such as the use of embedded links or drop-down menus) but involve web pages with little text on them (therefore requiring a low level of text processing).

The ERA tasks developed for PISA collectively populate most of the space in the imagined graph, with each task being classified according to the extent to which it demands both navigation and text processing.

To exemplify this conceptualisation of electronic reading, selected items from the 2009 Field Trial are described below. The items are part of a unit entitled “Let’s Speak”, and relate to an online discussion forum about the challenges of speaking in public. This set of stimulus is classified as *message-based* because of its potentially dynamic nature: visitors to the forum can alter its content by contributing their own responses.

The discussion forum has eight entries, which are presented on a single page. It is therefore necessary to scroll down the page to read the original entry. An embedded link within one of the posts in the discussion forum takes students to the only other piece of stimulus available within the unit – a page of “tips on public speaking”, recommended by one of the forum participants as “expert advice”.

Two of the four items administered as part of this unit in the Field Trial for PISA 2009 focus on main ideas (either within a post, or across the forum as a whole). These items require minimal navigation (scrolling only), but substantial text processing. They would therefore appear close to zero on the Y-axis but some distance along the X-axis of the graph in Figure 1.

The final item of the same unit asks students to consider all of the posts and evaluate their content in terms of the contributor’s professional credentials, or in terms of the quality and persuasiveness of the arguments presented. To respond, students construct their own post, which

then appears as part of the discussion forum. The text processing demands of this item (comparing and evaluating several short texts) are relatively heavy. In contrast with the items described previously, the navigations demands of this item are not simple: in addition to scrolling, students need to click on an embedded link, and then on another link in order to construct their reply, as well as using tabs and the back button. This item would appear some distance from the origin of both the X and Y axes in the graph in Figure 1.

As outlined above, the ERA instrument has been designed to systematically vary the weightings of text processing and navigation. This conceptualisation has also been used as the basis for a preliminary exploration of the features that condition item difficulty in electronic reading.

In our initial work, three researchers rated the ERA items. Ratings were given both in terms of variables related to text processing, and variables related to navigation. The variables related to text processing are collectively represented on the X axis of the graph in Figure 1. These variables were all rated on a scale of 1 to 4, in two groups: those related to the task (the interaction between the question and the text) and those related exclusively to features of the text. For rating the predicted difficulty of text processing, the three variables related to the task were:

- the type of retrieval/interpretation/ reflection, in terms of increasing complexity;
- the nature of competing information, in terms of its quantity and distracting power; and
- the nature of the information needed to answer the question, in terms of degree of abstractness.

Two text variables were also given ratings:

- familiarity of content and
- length/amount of text.

Ratings were also given to variables related to navigation. These variables are collectively represented on the Y axis of the graph in figure 1. In contrast to the text processing variables, the scale used to rate the navigation variables ranged from 0 to 4. This difference in scale reflects the fact that for some items there is no navigation involved at all – not even scrolling. Items of this kind involve interpreting a piece of text, all of which is visible when the question appears. A question of this kind would receive zero ratings on all the variables related to navigation. The

navigation variables that were rated were:

- the number of steps. The number of navigation steps required to complete the item ranges across the set of items in the ERA from 1 to more than 10.

- the amount of guidance. In some cases the item stem is very specific about the navigation steps that are required (for example, some items include an instruction to click on a specified link). The opposite extreme on this variable is represented by items in which the student must determine the series of navigation steps required with no guidance.

- the amount of distraction. This variable relates to what else is on the screen the student is seeing. As the number of (plausible) links and their attractiveness increases, ratings on this variable would also increase.

In order to explore whether this system of ratings is reliable, the correlation between the ratings of each pair of raters was examined. For the text processing variables, the correlation between pairs of raters was moderate, ranging between 0.70 and 0.78. For the navigation variables, the correlation between pairs of raters was noticeably higher, ranging between 0.84 and 0.85. The higher figure for the navigation variables is probably explained by the more quantitative nature of the task: the number of steps, for example, can be counted, whereas a text processing variable like type of retrieval/interpretation/reflection cannot.

The relationship between the predicted difficulty (the ratings) and the actual difficulty of a subset of the items in the Field Trial was also examined. The ratings described above were identified as the "predicted difficulty". The "actual difficulty" was based on the PISA 2009 Field Trial data, the results from 22 countries, with an average of 85 students per item from each language/country combination – about 2000 cases per item. As this was an exploratory study, only the first four units (comprising 27 items) administered in the Field Trial were rated by all three raters. The measure of the actual difficulty of each item was the percentage of students in the Field Trial (across all participating countries) who received full credit for the item. Each of the three raters had assigned eight ratings to each item (three related to the task, two related to the text and three related to navigation). For each rater, these ratings were added. An average rating for each item was calculated from these totals. The correlation between the average rating of the three raters and the percentage correct was then calculated.

For the text processing variables there was a modest correlation (0.63) between predicted and actual item difficulty. For the navigation variables, the correlation was lower (0.46). Neither of

these sets of variables alone, then, explains a large amount of variance in item difficulty. When the average rating for each item for text processing and navigation combined was calculated, however, the correlation rose to 0.72 – about 50% of the variance in item difficulty. That is, while text processing alone and navigation alone did not explain much of the variance in item difficulty, considering the two variables together increased the amount of variance in item difficulty that was explained.

This preliminary study provides evidence that both text processing and navigation contribute to the difficulty of items in PISA ERA, which is consistent with the framework and intention of PISA ERA. The work provides some support for the hypothesised differences between print and electronic reading and suggests that the ERA items do represent authentic features of reading in an electronic environment.

It is a commonplace that much of the growth in digital reading has occurred in the last few years. Between 2000 and 2007, Internet use doubled in OECD countries, from 33 to 68 people per 100, and more than trebled worldwide, from 7 to 22 per hundred (World Bank database, 2009). This trend leaves little doubt that electronic reading will become increasingly prevalent, and proficiency in electronic reading increasingly important for 21st century citizens. PISA ERA represents the first attempt in a large-scale international survey to assess the skills and knowledge required to read in the digital medium. In this paper we have described the way in which electronic reading has been conceptualised in the PISA 2009 framework and operationalised in the ERA instrument. The framework and instrument assume that both text processing and navigation are part of electronic reading. A preliminary exploration of factors that condition item difficulty in the electronic reading items offers some support for the hypothesis that both skills are integral to proficient electronic reading. This work is a first step in a process that will be continually refined as technological advances are made, and our understanding of the nature of reading in the electronic medium grows. Although the work is preliminary, this is one of the first studies that provides empirical evidence in support of reading in an electronic environment as qualitatively and quantitatively different in some important respects from print reading. This approach to understanding features that contribute to item difficulty has potential to inform the test development process not only for PISA ERA, but for assessments of electronic reading in other contexts. Ultimately it is educators and learners who

will benefit most from this work as our understanding of electronic reading deepens and we continue to give indicators of ways in which we can help students prepare to meet the demands of the information age.

References:

International Telecommunications Union. (2009). ICT statistics database. Retrieved 23 February 2009, from http://www.itu.int/ITU-D/icteye/Indicators/Indicators.aspx#

OECD. (2000). *Measuring Student Knowledge and Skills: The PISA 2000 assessment of Reading, Mathematical and Scientific Literacy*. Paris: OECD.

OECD. (2006). *Assessing Scientific, Reading and Mathematical Literacy: A framework for PISA 2006*. Paris: OECD.

OECD and STATCAN. (2005). *Learning a living: First results of the Adult Literacy and Life Skills Survey*. Paris and Ottawa: Organisation for Economic Co-operation and Development and Statistics Canada.

The World Bank. (2007). World Bank database. Retrieved July 10 2007, from http://devdata.worldbank.org/data-query/

World Bank database. (2009). The World Bank. Retrieved Retrieved July 28 2009, 2009, from http://devdata.worldbank.org/data-query/