A LOW PROFICIENCY TEST OF EXPEDITIOUS READING

Francisco Naranjo-Escobar
Lakeland University Japan (Japón)

Fan Chen
Nanjing University of Information Science and Technology (China)

Michelle Czajkowski
University of Melbourne (Australia)

Abstract: This project sought to develop a test to assess expeditious reading abilities in low proficiency English learners with Spanish L1, which to the date the project was conducted and to authors' best knowledge constituted a research gap. To achieve this aim, the team developed a new construct based on the existing literature and extrapolated from current research to fill the gap. Participants were Chilean high school students enrolled in the public education system. Results showed that low proficiency learners can perform expeditious reading operations in their L2. These results, however, must be taken with care as further research is required to reach conclusive results.

Keywords: expeditious reading, low proficiency, reading, language testing

Desarrollo de una prueba de comprensión de lectura expedita (expeditious reading) para estudiantes con bajo nivel de proeficiencia

Resumen: El presente estudio tenía como objetivo desarrollar un examen para medir la capacidad de aprendientes de inglés con bajo nivel y español como lengua materna para llevar a cabo operaciones de lectura expedita —expeditious reading. Este tipo de habilidad, a la fecha de desarrollo del proyecto y de acuerdo al conocimiento de los miembros del grupo, corresponde a un vacío investigativo que merece atención. Para lograr dicho objetivo, el equipo desarrolló un nuevo constructo basado en la literatura relevante existente y extrapoló los resultados existentes a la fecha para llenar el vacío investigativo identificado. Los participantes fueron estudiantes de educación media chilenos que asistían al sistema público. Los resultados sugieren que los aprendientes con bajo nivel pueden llevar a cabo operaciones de lectura rápida en su segunda lengua; sin embargo, se requiere mayor investigación para demostrar que los resultados son generalizables a toda la población.

Palabras Clave: lectura expedita (expeditious reading), bajo nivel de competencia, lectura, desarrollo de exámenes estandarizados

* Para correspondencia, dirigirse a: Francisco Naranjo-Escobar (naranjo@japan.lakeland.edu); Fan Chen (850222@nuist.edu.cn) y Michelle Czajkowski (michelle.czajkowski@unimelb.edu.au)

1. Introduction

In recent years, the development of L2 reading test has placed greater attention on expeditious reading, which has been neglected in traditional L2 teaching and testing (Vidakovic, & Dimitrova-Galaczi, 2013). This shift in attention was in part motivated by the wide recognition of Weir and Khalifa's reading model (2008), which not only features a taxonomic framework of reading types but also links those types to the underlying mental processes. Equipped with this model of reading, several validation practices have been conducted to investigate how well expeditious reading is addressed in English language tests (e.g., Khalifa & Weir, 2009; Weir *et al*., 2009a; Weir *et al*., 2009b; Weir, Huizhong & Yan, 2000). These studies show that expeditious reading —namely the ability to read a high volume of text during a short period— is not tested independently or addressed fully in some academic English tests due to practical issues in test development (Weir *et al*., 2009a). General English tests, however, tend to place even less emphasis on expeditious reading skills, probably due to the absence of such skills in CEFR descriptors at lower proficiency levels. This absence has raised our concern, as in theory, expeditious reading is not cognitively more demanding than other types of reading (Weir & Khalifa, 2008) and thus could be attempted by lower-level learners. To address the issue raised above, the present study developed a test to measure expeditious reading skills of lower-level English learners whose test results could be used to inform relevant pedagogical decisions.

The stages of test design were characterized as construct identification, prototyping, pilot testing, and field testing (Enright *et al*., 2000). A brief description of each of these stages follows:

1) Construct identification: the construct was described based on a cognitive processing model of reading, and possible operationalizations of the construct were proposed;

2) Prototyping: possible tasks were prototyped and scoring schemes developed under the guidance of draft test specifications;

3) Piloting: two rounds of piloting were conducted in timed conditions to improve on prototyped tasks, with quantitative results yielded to establish test reliability and qualitative data elicited from a retrospective survey to add to validity evidence.

This paper will first discuss construct identification in a literature review, followed by a methodology section to describe prototyping and piloting procedures. The results of the prototyping and the piloting processes will be analyzed to provide the validity argument for our test.

2. LITERATURE REVIEW

2.1. Construct Identification

The aim of this literature review is to identify the construct of our reading test. Essential topics that were investigated include the following:

1) A literature review of theoretical concepts of expeditious reading and empirical studies of how this type of reading differs from careful reading, which provided a basis for describing the construct;

2) a survey of literature on how expeditious reading is tested in previous and existing tests, which helped to further identify the construct in the target domains;

3) a survey of relevant reading papers to further specify the target domain on which the test prototyping can be based.

2.2. Conceptualizing reading

Reading viewed as "a complex cognitive process the reader is engaged in with a written text for one or more specific goals" (Shiotsu, 2010), has been studied within the taxonomy of skills and subskills. Amidst the most influential framework of reading, Carver (1992) differentiates among five types of reading: "scanning", "skimming", "rauding", "learning", and "memorizing" by reading goal, rate, and mental processes activated. Alternatively, Grabe and Stoller (2002) distinguish between seven reading types, ranging from "reading to skim quickly" to "reading to integrate information". A similarity shared by the two models is that they both make a distinction between slow and fast reading rates, with "scanning" and "skimming" representative of fast reading and "reading to learn" of slow reading. Also, the cognitive processes typical of different types of reading are likewise specified in both models. For instance, Carver (1992) suggests that "scanning" and "skimming" are two basic types of reading, which only involve lexical access and semantic decoding in terms of the culminating mental processes, whereas "learning" and "memorizing" require higher-order processes such as "idea remembering" and "factual rehearsal".

While the two models take careful consideration of reading types and link each type to relevant components of cognitive processes, the description of the context in which the cognitive processes are involved is scarce. To fully capture the nature of reading, this study selects Weir and Khalifa's (2008, p. 4) cognitive processing model (see Appendix A for illustration), which features a more systematic distinction of reading activities (Urquhart & Weir, 1998) as well as an adequate emphasis on the relevant "cognitive processes mediated by the contextual parameters of the text and task" (Khalifa & Weir, 2009).

Firstly, reading types in this model are classified into "careful" and "expeditious" while reading levels into "global" and "local". This taxonomy is considered an

improvement on previous models due to its "dynamic" representation of reading modes in a full range (Moore, Morton & Price, 2007, p. 9), as opposed to a list of discrete types. According to Urquhart and Weir (1998), careful reading and expeditious reading differ in selectivity: the former involves examining the whole text submissively, whereas the latter involves a deliberate selection of the parts to read. In turn, when readers skim the text (i.e. expeditious reading global), they tend to process the whole passage quickly and locate critical information at the introductory and concluding paragraphs. Furthermore, when scanning the text for specific information (i.e. expeditious reading local), readers will only focus on certain parts of the passage to find the match. A third operation is search reading, which can be conducted at either global or local level and involves deeper processing of a predetermined topic.

Further, when the types of reading are linked to the components of cognitive processes, a cline of cognitive demand can be detected among the four reading types, ranging from scanning to careful global reading. The level of cognitive demand, according to Weir and Khalifa (2008, p. 9), is as follows in ascending order of difficulty: scanning/searching for local information, careful local reading, skimming for gist, careful global reading for comprehending main idea(s), search reading for global information, careful global reading to comprehend texts, and careful global reading to comprehend texts.

Scanning is the least demanding because it usually requires word recognition only. Processing up to the sentence level is typical of careful local reading. Text level representation is the goal of both skimming and careful global reading, with the latter exploiting more cognitive resources to create a more detailed macrostructure. Interestingly, skimming sometimes only involves sentential processing when the text and genre knowledge lend adequate support to build a macrostructure. The difficulty of search reading —on the other hand— varies greatly depending on the level at which it is conducted. The hypothetical order of difficulty across the reading types is supported empirically with a study into the First Certificate in English (FCE) and the Cambridge English: Advanced (CAE) items (Rose, 2006).

Weir and Khalifa's (2008) hypothesized order in cognitive demand, however, does not imply any relative importance among the reading types. Each type assumes equal status in this model, and the choice among them depends on reading purposes and demands. Considering the changing demands that readers may encounter in real-world situations, different types of reading should receive adequate coverage when defining the constructs of a reading test (Alderson, 2000). However, studies show a preponderance of careful reading over expeditious reading in language testing, probably because of pedagogical influences and practicality concerns (Weir *et al*., 2013). In fact, the cognitive processing model was proposed in response to the bias towards careful reading in language testing and aimed to give expeditious reading due attention in examining and operationalizing the construct of reading. Our study is motivated by this observation that the status of expeditious reading should improve in test development. Therefore, in an attempt to address the potential neglect of this reading type in language testing, we carried out a preliminary survey of existing

reading papers and relevant validation reports to identify to what extent and in what way expeditious reading is under-represented in language tests.

2.3. Testing Expeditious reading

As mentioned above, careful reading has attracted overriding attention over expeditious reading in language testing, especially in the U.K. Part of the reason lies with British pedagogical traditions, where texts were viewed as means of teaching vocabulary and grammar instead of reading skills. Time constraints in the classroom teaching also rendered the processing of long texts less practical (Weir *et al*., 2013).

Recent decades, however, have witnessed a surge of interest in featuring expeditious reading in teaching materials (e.g. Grellet, 1981; Paran, 1991) probably due to the emerging recognition that international students experience difficulty in coping with large volumes of reading at English-medium schools and universities (Weir, 1983; Robb & Susser, 1989, Weir *et al*., 2009a). Test developers have also begun to encourage expeditious reading skills by using particular task types or imposing time constraints.

For example, the TOEFL iBT reading test features "reading to find information" items, which are intended to encourage the use of skimming and scanning skills (Educational Testing Service, 2012). Alternatively, "matching headings" items in IELTS reading modules are designed to target expeditious reading skills, with the official guide (Cullen, French & Jakeman, 2014) featuring reading strategy tutorials that demonstrate how skimming and scanning can be used to complete tasks. Despite the effort to operationalize expeditious reading using particular item types, validation reports for the two tests in question reveal that task types cannot prove a good predictor of reading strategies, although some particular items do encourage the use of expeditious reading according to self-reports by some examinees (Cohen & Upton, 2007; Weir *et al*., 2009a; Weir *et al*. 2009b). Weir *et al*. (2009b) further suggest that, albeit practically difficult, time limits should be imposed on IELTS to enforce the use of expeditious reading skills.

In China, however, Test for English Majors (TEM), College English Test (CET) and Academic English Reading Test (AERT) all have a separately timed section devoted to testing expeditious reading (or "fast reading"). Research on these tests has yielded empirical evidence that candidates perform differentially on two reading types (careful versus expeditious). Retrospective data further suggests that poor performance on AERT reading test results from unawareness of different reading skills and poor linguistic proficiency might pose "a threshold for the effectiveness of the reading skills and strategies" (Weir *et al*., 2000, p. 134).

Perhaps due to the perceived "threshold" in the effective use of reading skills, expeditious reading features less significantly in English tests aimed at lower-level learners. In validating Cambridge Main suite reading papers—namely Cambridge English: Key (KET), Cambridge English: Preliminary (PET), Cambridge English: First (FCE), Cambridge English: Advanced (CAE), and Cambridge English: Proficiency (CPE), Khalifa and Weir (2009) suggest that expeditious reading is addressed in

all levels except in the case of the A2-aligned KET, where the test focus is placed on careful reading solely. The test design in terms of representation of expeditious reading can be explained by referring to the ability descriptors across CEFR levels. For general reading ability, A2-level learners are expected to deal with short, simple texts on familiar topics, while B1-level learners can scan longer texts. In addition to considerations of text difficulty, the lack of expeditious reading at the A2 level is further justified regarding reading flexibility (Khalifa & Weir, 2009, p. 68). Under the "reading for orientation" description, the ability to scan is expected from B1 and above, with the reading flexibility expected to increase with proficiency level.

The specifications at each level, however, are still arguably vague and limited: there is no specific demand of reading speed imposed on each level, and the suggestions as to text length are also open to interpretation (Khalifa & Weir, 2009). Therefore, while lower-level learners may not flexibly adjust reading speed and strategy to their reading purposes, it seems hasty to conclude that they cannot expeditiously read simple texts of reasonable length when given sufficient instruction. This assumption can be backed up with several inferences drawn from cognitive processing models, ALTE can-do lists, and L1/L2 reading research. First, according to Weir and Khalifa's (2008, p. 9) hypothesized order of cognitive difficulty among reading types, A2 learners who can read carefully at a local level can also do scanning as scanning is the least cognitively demanding task within their hierarchy. Such learners are also expected to do skimming when the familiarity of the genre helps reduce the cognitive processes to the sentence level. Second, the text types suggested at the A2 level include sections in newspapers, magazines, encyclopedias, which are usually approached in expeditious reading mode. Third, a number of empirical studies show evidence that L1 reading strategies may transfer to L2 reading (Jimenez *et al.*, 1996; Hua, 1997, as cited in Koda, 2005). It can thus be hypothesized that those who can read expeditiously in their L1 can also do so in an L2.

Given the assumption of A2 learners' ability to read expeditiously, although not explicitly stated in either CEFR descriptors or the ALTE can-do list, one may argue for a potential construct underrepresentation in the KET reading paper. With expeditious reading neglected in the test, English courses and materials targeting the A2-level might also place less focus on the development of such skills, which are essential in improving reading efficiency and accuracy. Students may also find it challenging to gain reading facility with limited access to relevant instruction. The lack of assessment tools of expeditious reading may also present a challenge for teachers who need a reliable measure of such abilities to inform instructional decisions and scaffolding strategies. The concern for the potential neglect of expeditious reading in KET papers is compounded by the fact that the B1-targeted PET reading paper features both skimming and scanning items in its longest texts (550 words as opposed to 250 in KET), while the B2-targeted FCE reading paper has search reading items, which tap into more extensive cognitive processes. With the abrupt rise in the reading demands from A2 to B1-above tests, elementary learners might feel underprepared to cope with the intermediate level requirements; furthermore, since instruction on

expeditious reading does not start at beginner levels, test-takers might feel even more underprepared.

In an attempt to address the potential issue of under-representation of expeditious reading in CEFR and relevant tests at the A2 level, this study aims to develop a test to measure the ability of expeditious reading of lower-level English learners by imposing time control on reading tasks. Since there is no existing standard on the extent to which these learners can comprehend texts expeditiously, this test is designed as a preliminary instrument to discriminate between learners' ability of expeditious reading without relating their performance to any criterion; test results are expected to assist in the tentative efforts to draft assessment criteria of expeditious reading. Further, considering that search reading involves more cognitive processing and that it is not introduced into Mainsuite papers until the B2 level, as well as that it may create greater difficulty for lower-level learners, the construct of expeditious reading in our test is restricted to skimming at a global level and scanning at a local level (see Table 1 for a definition and analysis of the construct).

Expeditious reading		
Sub-types of reading	**Skimming**	**Scanning**
Purpose	Processing a text selectively to get the main idea(s) and the discourse topic as efficiently as possible,	Looking quickly through a text, not necessarily following the linearity of the text, to locate a specific symbol or a group of symbols
Operationalizations	As appropriate to text-type: -reading title and subtitle quickly; -reading introductory and concluding paragraphs carefully; -reading first and last sentence of each paragraph carefully -glancing at words or phrases	Looking for (matching): -specific words/phrases; -figure percentages; -dates of particular events; -specific items in index; -names
Focus	Both global and local	Local
Text coverage	Selective reading to establish important propositions of a text	Ignoring most of the text
Rate of reading	Rapid with some careful reading	Rapid with some careful reading

Components of cognitive processing involved	Decoding – lexical access- syntactic parsing – establishing propositional meaning – inferencing – building a mental model – (creating a text-level representation)	Decoding – lexical access- (syntactic parsing)
Knowledge base	Lexical, syntactic, genre, (topic)	Lexical, (syntactic)

(Source: Urquhart & Weir, 1998)

Table 1. *Definition and analysis of the construct*

3. Contextualization of this study

With the test construct specified in terms of its operations and cognitive processes, the study moves on to the next stage of identifying features that are representative of the target domain, i.e., reading tasks geared at lower level learners. This stage entails identifying the range of text features characteristic of A2-B1 levels and assessing the contribution of various features to text difficulty. This was carried out by collecting sample test materials available on the Cambridge ESOL website and other publications. Due to the limited availability of relevant resources, the scope of the survey was restricted to KET and PET examination resources, such as sample reading papers, examination handbooks, relevant validation reports, and test preparation materials. The survey results are listed in Appendix B (Task features of relevant reading papers), which provide contextual parameters, such as text lengths, genres and test techniques for developing the test specifications. Time constraint for expeditious reading is not particularly specified in either paper, and a survey of literature only shows an estimated reading speed for B2-above-learners (Weir *et al*., 2000). Therefore, trialling research is needed to identify the appropriate time constraint to enforce the use of expeditious reading.

In summary, the goal of our study is to develop an expeditious reading test targeted at lower level English learners. The literature review has described the theoretical framework on which the construct is identified and further proposed guidelines for operationalization of the construct. The following processes of test design involve a prototyping phase to model test specifications and two piloting phases to analyze the suitability of the test tasks and representation of the construct. To collect validity evidence in each phase, four research questions are formulated to guide the following development activities:

1) What is the appropriate time constraint to enforce the use of expeditious reading in our test?

2) Will candidates of the test use the set time fully?

3) To what extent will this test differentiate between overall proficiency levels?

4) Will the candidates report the use of expeditious reading strategies?

4. METHODOLOGY

The design of this test comprised three different stages which had their own methodological constraints. Each of them has been exposed separately, albeit there are shared general aspects. This section starts by explaining methodological aspects shared by all the stages of this project, outlined in Table 2.

Stage	Stage identification	Stage purpose
1	Prototyping (time allocation pilot)	Check ideal time for test completion
		Item difficulty and item discrimination analyses with results
		Check overall reliability
2	Final Form (first timed pilot)	Pilot of test with timed conditions
		Correction of eventually problematic items detected during stage 1 (prototyping)
		Item difficulty and item discrimination analyses with results
		Check overall reliability
3	Final Form (second timed pilot)	Item difficulty and item discriminations analyses with results
		Check overall reliability
		Get data to contrast first-timed and second-timed administrations.

Table 2 *Test development project stages*

4.1. Participants

This test is being designed for a wide audience comprising learners from a variety of L1s. However, the accessed sample was three groups of native speakers of Spanish whose ages were between 13 and 15 years old. Each group received one letter for identification (A, B, and C) and took the test once to avoid training effect (Brown, 1988). Participants were high school students from Viña del Mar in Chile and were contacted via a proxy working for the institution.

Proficiency ranged from A1 to B1 according to the teacher's assessment, and the three groups had a similar number of participants (group A: 27; group B: 33, and group C: 24). Additionally, the proxy shared student's grades (based on the Chilean assessment scale as per the Ministerio de Educación (1999) standards).

To ensure anonymity and tracking of results, each student received a unique alphanumeric code showing its group membership and roll call order. Therefore, the first participant for every group was: A1, B1 and C1, respectively. Group A participated

in the prototype administration (time allocation pilot), whereas groups B and C took the Final Form during the first- and second-timed administrations.

Participants were not informed about the main objective of this project; nevertheless, they were debriefed once they sat for the test by the proctor.

Participation in the project was granted via verbal acknowledgement. This was given by the school principal under the condition that the results be shared with the school after the project finished.

4.2. Platform for test administration

Due to the geographical distance and time differences between Chile and the location of the team, the research team chose to administer the instruments electronically. The selected platform was the JotForm website [www.jotform.com]. It allowed us to accurately measure the time candidates spent during the time allocation piloting without human intervention. Likewise, it allowed randomization of question display in-test.

4.3. Instruments layout and visuals

All instrument versions used the same visual layout and had six electronic pages. Instructions on how to answer the test were written in Spanish under the assumption that attention is a limited resource (Schmidt, 2001). This was done to prevent cognitive resources from being devoted to decoding instructional content.

Questions were presented before the texts to provide students with an objective for reading.

4.4. Instruments

There were two instruments, a *Prototype* used for time-allocation and a *Final Form* for two separate administrations. The Prototype was coded as Form A, while the Final Form was called Form B.

4.5. Prototype (time-allocation pilot)

The Prototype —Form A— featured a mechanism to measure the time students spent on the test. This was possible via a script available on the JotForm website [www. jotform.com], which allowed the calculation of an average time spent by participants.

To ensure a varied difficulty, all texts were subjected to a Flesch-Kincaid readability scale analysis in MS-Word, as shown in Table 3.

Text	Flesch easiness scale
Text I	88
Text II	72
Text III	77

Table 3. *Readability scale of texts for the prototype*

4.6. Question types

Based on Weir's (1993) recommendations, all items have one unequivocal answer, are worded with vocabulary that is not harder than the one found in the readings, and avoid the use of mathematical operations. To cater for the last recommendation, all numbers were converted into word notation to prevent L1 interference.

There were 17 questions in total of which eight were Short Answer Question items, and nine were multiple-choice items. This phase was primarily designed to calibrate the ideal time to articulate expeditious-reading operations and detect potentially problematic items.

Question types were as follows:

Multiple choice items (MC).

Items in this category targeted skimming and scanning reading abilities. These items have three distracters and one answer. Only during the time-allocation pilot was there an extra option (*I don't know/No sé la respuesta*) to prevent candidates from randomly selecting answers to end the test fast. This forced test-takers to engage in reading the questions and texts. Each question was weighted 1 point out of the total score. Wrong answers did not affect the total score.

Short Answer Question items (SAQ).

This type of item constituted a fill-in-the-gap question created to evaluate scanning. Although the design required candidates to input linguistic data, there is no danger of construct irrelevant variance caused by an increase of difficulty (Urquhart & Weir, 1998). This has been prevented by requiring candidates only to transfer specific information from texts and not to produce it based on any read contents.

While the literature discourages the use of gap-fill SAQs because of its low cognitive demand (Haladyna & Rodriguez, 2013), we have mitigated that drawback by adding extra cognitive load through the enforcement of a time limitation.

Each item was equally weighed as 1 point of the total scale except for two items in the time-allocation pilot (items 9 and 10) that had a 0.5-point weighing and featured a double gap.

4.7. Final Form (timed pilots)

The Final Form, used during the first and second timed pilots and referred to as Form B, featured the same question type and texts as the prototype. However, it had 15 questions in total (weighed 1 point each), of which six were SAQ items and nine were MC items. This decision was made based on the results from the prototype administration (see the Results section), which brought about the change of SAQ items from the double- to single-gap format.

This form had a 20-minute time limit derived from the prototype administration results (see the Results section).

Texts for this stage, as above mentioned, were the same as those for the prototyping stage but with minor changes based on the prototype administration results (see the Results section). Table 4 shows the Flesch-Kincaid readability scale for the texts for this form.

Text	Flesch easiness scale
Text I	88
Text II	75
Text III	80

Table 4. *Readability scale of texts for the Form B*

The Final Form (Form B) had a short strategies survey used to investigate whether candidates engaged in expeditious- or extensive-reading operations. However, the survey was only implemented for Group C (October 15 administration) and featured ten questions in total. Questions constituted statements regarding candidates' behaviour, and respondents had to select from three options: *Sí/Yes*, *No/No*, and *No sé/I don't know.* The last option was added to control for variance should candidates not know what to answer. Instructions were given in Spanish with a visual example on how to answer the survey. Response collection was automatic and electronic via GoogleForms.

4.8. Procedures

Administration of the prototype

Administration was successfully conducted on Tuesday, October 6, 2015, from 12:43 to 13:10 Chile time (GMT-3). This stage provided the time allocation for the subsequent administration and shed some light on items requiring attention. Based on the results from this administration (see Results section), the following changes were made:

a.	SAQs items were modified to have only one gap, with each gap weighed at 1 point.

b.	Text II was modified so that the environment surrounding the location of the answer for the modified SAQ items satisfied the new single-gap format.

c.	Text III was simplified to improve the response rate (passive sentences were simplified).

On the other hand, and regarding the ideal time calibration, this stage yielded a final time allowance of 20 minutes (see the Results sections for further details)

Administration of the Final Form (Group B).

The test was successfully administered on October 9, 2015, between 12:43 PM and 13:10 PM Chilean time (GMT-3). Based on the results from this administration (see the Results section for further details), it was decided that the test form employed during this administration would be kept for the next one.

Administration of the Final Form (Group C).

The test was successfully administered on October 15, 2015 between 09:02 AM and 09:22 AM Chilean time (GMT-3).

Analyses for all administrations.

These statistical procedures were run for the data obtained during all piloting stages.

Descriptive statistics: S.D. and mean; item difficulty analysis**:** to investigate how easy the items were for test-takers; item discrimination analysis (multiple regression): to investigate to what extent test items discriminated among candidates' ability to perform expeditious-reading operations; Pearson correlation between overall test score and GPA scores; Cronbach's α: to investigate overall test questions reliability; a one way ANOVA: to investigate differences in overall score in test Form B and CEFR level. For this analysis, Groups B & C were pooled together as both took Form B; finally, a two-way ANOVA: to investigate possible interaction effects between test forms (A and B) and students' CEFR level. For this analysis, groups B & C were pooled together (Form B) while group A (Form A) remained independent.

5. RESULTS

Descriptive Statistics

Three test sessions (groups A, B and C) were run using two different test forms (Form A, B), each with three texts/sections (text I, II, III). Each section is worth 5 points, and the total maximum score for the test is 15.

Table 5 summarizes participants and test versions for the three sessions.

Group	Form	Date	N	A1	A2	B1	GPA	(SD)
Group A	Form A	October 6	25	19	5	1	4.59	*(0.76)*
Group B	Form B	October 9	33	19	9	5	5.26	*(.79)*
Group C	Form B	October 15	24	8	11	5	5.88	*(.50)*

Table 5. *Group, Form and Participant Numbers and CEFR Levels*

Prototype

This test form (Form A) was used with Group A to determine time and to identify possibly unsuitable items. Mean time for this group was 19.95 minutes (SD = 5.81), and based on this, a time limit of 20 minutes was imposed on Form B, the pilot version. This test form was found to have high reliability as measured by Cronbach's alpha (α = .804). However, facility was quite low for many items, with an average facility index of 0.32, which is reflected in the test results, M = 4.76 (of 15), SD = 3.14. Problematic items from Form A were analyzed and modifications were made to stems, keys and distractors in response.

A summary of item analysis statistics can be found in Appendix C.

Pilot: Group B

Form B, the revised and final form test, was piloted with a second, independent group of students. Form B was a timed test at 20 minutes, based on the mean time for Group A. For this group, N = 33, the mean score was 7.09 (max. 15) (SD = 3.26), and the mean time taken was 18.61 minutes (SD = 1.95), with a minimum of 12.38 minutes. Reliability for this Form was slightly lower, α = .768, but mean item facility had risen (.047), with most questions in an acceptable range. No further changes to the test were made, and Group C was given the same test.

A summary of facility and discrimination index figures can be found in Appendix C.

Pilot: Group B and C combined

As both Group B and Group C took Form B, these two groups were combined for a final analysis, summarized in Table 6.

	All (N=57)		A1 (N=27)		A2 (N=20)		B1 (N=10)	
	M	**(SD)**	**M**	**(SD)**	**M**	**(SD)**	**M**	**(SD)**
GPA	5.52	*0.74*	5.10	*0.72*	5.75	*0.52*	6.21	*0.42*
Text I	3.11	*1.41*	2.59	*1.39*	3.35	*1.39*	4.00	*0.94*
Text II	2.74	*1.52*	1.67	*1.04*	3.35	*1.14*	4.40	*1.07*
Text III	1.67	*1.43*	1.19	*1.36*	1.80	*1.20*	2.70	*1.57*

| Total | 7.51 | 3.38 | 5.44 | 2.56 | 8.50 | 2.56 | 11.10 | 3.00 |
| Time | 18.42 | 2.33 | 18.93 | 1.77 | 18.24 | 2.17 | 17.40 | 3.55 |

Table 6. *Descriptive Statistics for Group B+C*

Use of time limit

The maximum time allowed, 20 minutes, was used by the majority of test-takers, and there is no statistical difference between mean time use and CEFR level, $F(2, 54) = 1.719$, $p = .189$, $\eta^2 = .059$

Student level discrimination

A one-way ANOVA, used to investigate the differences in mean scores between CEFR levels A1, A2 and B1, showed a significant difference between the mean total scores of each level, $F(2, 54) = 18.91$, $p<.001$, $\eta^2 = .411$. To look at this association from a different angle, a Pearson correlation showed a moderate correlation between GPA and total test score, $r = .331$, $p = .012$.

Skimming and scanning

A Pearson correlation investigating the relationship between performance on skimming items and scanning items showed a strong positive correlation, $r = .483$, $p<.001$. The mean of the weighted[1] skimming score was compared with that of the scanning score in a repeat measures t-test, which showed no difference between the two sets of sub-skill questions, $t = .123$, $p = .903$.

Text item analysis

Cronbach's alpha for the group was found to be .770 (15 items), a result comfortably in the acceptable range for pilot tests.

Item analysis (Table 7) shows the Facility Value and Discrimination Index for each item. Most items fell within a desirable range of .33-.67 for facility and all but two above the .3 threshold for discrimination.

Item	Facility	Discrimination
1	0.58	0.33
2	0.77	0.23

---

[1]   The score (max 3) was multiplied by 4 to achieve a maximum score of 12, comparable to the scanning items with a max of 12.

3	0.49	0.40
4	0.65	0.53
5	0.61	0.67
6	0.67	0.57
7	0.51	0.48
8	0.30	0.63
9	0.91	0.32
10	0.35	0.74
11	0.44	0.32
12	0.44	0.60
13	0.40	0.50
14	0.26	0.28
15	0.12	0.67

Table 7. *Group B+C: Item Analysis, (N = 57)*

Retrospective survey on reading skills

Group C (N = 24) was asked to complete a survey on reading skills directly after the tests. The survey had ten questions, three negatively scored, regarding expeditious reading. Many replies (53 %) reported positive use of expeditious reading skills, though individual scores show a range of 0 to 10 (M = 5.33, SD = 2.50). Replies indicate that looking for specific words and reading the questions first were the two most used strategies, and most participants reported reading only relevant sections. However, replies indicate that prediction was not widely used, and while many report only reading selected sections, most report trying to read everything.

Analysis: Form A and Form B

To investigate whether the modifications made in Form B produced significant changes in results, a two-way ANOVA was run to detect any interaction between test form and CEFR levels A1 and A2 (B1 was removed from this analysis because there was only one B1 who took Form A. No such interaction effect was found (F(1,57) = .020, p = .887.).

6. Discussion

The aim of this project was to design a test to indicate how well students at lower levels of proficiency performed at expeditious reading skills. This norm-referenced test was benchmarked against the lower CEFR levels for reading ability and hoped to

help identify those students who, at these lower levels, may lack these reading skills that start to become emphasized at intermediate levels.

The pilot test designed for this study aimed to measure low-level students' ability to use expeditious reading. We analyzed a number of features of our test to investigate the degree to which this instrument measures the target construct:

- Use of time limit across student proficiency levels

- Total test score across student proficiency levels

- Retrospective survey on reading skills used.

There was no significant difference between the use of time and the different CEFR levels of the participants in the pilot, showing that regardless of overall proficiency level, most participants made use of the total time. In addition, total test scores were shown to be significantly different across CEFR levels, with B1 achieving, on average, a higher score than A2, who achieved higher than A1. The test itself showed high reliability for a pilot, and items fell into an acceptable range in facility and discrimination. These results in themselves do not show that scores reflect expeditious reading skills. To investigate more thoroughly the degree to which expeditious reading skills are being tested, we can isolate different features of the test and ask a series of questions to test our inferences.

Are the participants reading expeditiously?

In the literature review, we focused on the KET (A2) and PET (B1) as examples of reading tests that focused on the ability of candidates at a certain level and used these tests as an anchor for comparison because the corresponding CEFR level assumed test-takers' capabilities. To investigate the degree to which our participants may have used careful reading instead of expeditious reading to achieve higher scores at higher proficiency levels, we can compare the text load in terms of total words divided by total time. For KET (A2), this ratio ranges from 18.5 to 20 words per minute. For PET (B1), the ratio rises slightly to 29 to 32. Both tests focus mainly on careful reading, with some expeditious reading introduced in the PET (Appendix B). In our pilot test, this ratio was 73.75. This difference in WPM ratio recalls the reading rates for different purposes outlined in Carver (1992), discussed in the literature review, where skimming and scanning are estimated to be done twice as fast as more linear and comprehensive forms of reading. The comparatively high ratio for our test strongly indicates that students at these two levels, A2 and B1, would not have the automaticity and reading speed to deal with the texts in our pilot without using expeditious reading skills.

However, this inference would be compromised if the texts in our test were easier to read as compared to our anchor tests. While KET and PET specify texts of a Flesch-Kincaid Reading Ease score of 78.3 and 64.7, our texts average was 81. This may suggest that the texts chosen were more suitable in terms of reading ease for A2 participants, and the relative ease of the texts for B1 participants may have allowed them to read more in a careful rather than expeditious manner.

Finally, the participants themselves generally reported using a range of expeditious reading skills when attempting to answer the questions. Whether or not this was due to their ability to determine such skills as appropriate ways to obtain the needed information or a result of the instructions which explicitly instruct them to use such skills may, from one angle, be irrelevant, as we are not testing their ability to determine which reading strategy might be appropriate but just their ability to employ expeditious reading skills.

However, it is worth considering the effect that such explicit instruction may have had on survey results; self-report can show different results if participants answer in a way they feel is expected by the researchers.

Overall, there is strong evidence suggesting that students at these levels would not be equipped to complete this test using careful reading only, an inference backed by the students' own introspection into their cognitive processing.

Do differences in participant test scores indicate differing ability in expeditious reading?

We have shown that the participants would likely need to employ expeditious reading skills when attempting this test. However, this reasoning also needs support showing that the participants could reach a degree of success in answering questions correctly.

Overall performance in terms of total score shows a general inability to perform expeditious reading at the A1 level. Participants at this level only achieved a mean score of 36 %. However, A2 and B1 level participants showed much better performance, 56 % and 74 %, respectively. While our test is norm-referenced, it does reflect the broader CEFR descriptors, and as such, it is worth noting that a passing score for KET and PET, our two anchor tests, is 60 %, and so despite the demonstrable differences between texts in these two tests and our pilot, the results from our participants show comparable success.

As discussed above, these scores in and of themselves do not necessarily indicate a better ability to use expeditious reading skills. However, the similar average reading ease score, nearly identical to the KET (A2), suggests that scores from participants are not artificially affected by text difficulty. This then leaves text length as the variable likely to be influencing scores the most, which, if true, indicates that expeditious reading is likely to explain the differences in scores.

For which level is the test best suited?

A1 participants averaged relatively low scores, and B1 participants performed well enough to show a ceiling effect. However, A2-level participants showed an even and wide distribution of scores. As text difficulty most closely aligns with the A2-referenced KET in terms of ease of reading, our test seems most suited for A2-level learners.

Are participant test scores generalizable to a broader context?

In the sections above, we have shown why it is likely that expeditious reading, as defined in our literature review, is the main variable being tested in our pilot. What is not yet clear, however, is the degree to which test results mean success in similar tasks in real life. We discuss this in terms of cognitive skills and context coverage.

One key difference between our test and many other tests that include expeditious reading in the construct is the fact that the purpose of our test, and the skills required, are introduced to the test-takers in the instructions, and so not only do participants know what information they are looking for (the test items), they know how they are meant to look. This is different from tests that include both expeditious and careful reading, where the candidate is meant to determine which skill is appropriate. In Weir and Khalifa's (2008) cognitive model of reading (Appendix A), these meta-cognitive skills, referred to as the Goal Setter and the Monitor, represent reading flexibility, and are fundamental aspects of skilled reading in a real-world context. The degree to which test preparation pre-informs the Goal Setter and the Monitor is certainly not negligible, but this distinction will speak to the generalizability of our test; high scoring participants may have shown an ability to employ expeditious reading skills but have not demonstrated the ability to determine that these are the appropriate skills to use in this given context. This is an important distinction, as success in reading exams, and by extension, reading in general, is likely linked to knowledge of metacognitive strategies as much as in reading skills (Bax, 2013).

Another limitation to this inference comes from the limited nature of the texts chosen. Successful test takers have demonstrated their ability to handle the test items in the three texts chosen, but although we have tried to align these texts with CEFR guidelines, they do not, in themselves, represent the breadth of the range of possible texts. The pilot test was developed under the guidance of draft test specifications (Appendix D), yet further piloting with different test versions will be required to show generalizability to the wider set of text types and genres identified as appropriate.

Finally, our pilot was limited to two item types: multiple choice and short answer questions. We chose two types of items to reduce a potential source of bias, but the degree to which high scores on these question types speak to real-world ability has not yet been determined.

Limitations and continued test development

The section above outlines two limitations that weaken the link between test scores and real-world ability. The first refers to the limited range of text types tested, and the second to the limited range of item types. Both issues can be addressed with further testing, which would include the development of test versions with different texts and items based on the test specifications, followed by comparisons between groups.

Other limitations which would need to be addressed in future development phases include the participant demographics, test section timing and the delivery method. Firstly, our participants now have included only a limited demographic. The degree to which their results can be generalized to a broader population is yet unclear. As discussed in the literature review, the ability of lower-level students to engage in expeditious reading skills may depend on L1 reading ability, which would be a function not only of individual differences but of educational setting and L1. We cannot at this point infer similar results in different educational and L1 settings.

Secondly, while we could determine and set a seemingly appropriate time limit, we do not have information on the time that participants spent on each of the three texts, a deficiency of many reading tests (Alderson, 2000), which can be best addressed with timed sections (Weir, 2005), particularly if expeditious reading is to be activated for the task.

Thirdly, the online delivery method may introduce sources of error that are construct-irrelevant. Test takers are required to scroll up and down to navigate between the text and the questions, which may increase demand on working memory compared to a paper-based test where the questions and the texts are presented side by side (Sanchez & Wiley, 2009), though this scrolling may, on the other hand, aid in skimming and scanning (Bernard, Baker, & Fernandez, 2002). This is one area where there is limited research, particularly in regards to L2 reading. A second issue related to online delivery is the fact that test-takers often copied and pasted answers for the short answer questions, a practice that may be eliminating a source of error in traditional delivery modes, e.g. orthographical errors, but may pose different issues in areas such as the development of scoring rubrics.

Finally, future piloting must control for student level more rigorously. We were given estimated CEFR levels for our participants by their teachers and were provided GPAs as a point of comparison. While we have no reason to suspect that the CEFR levels provided were inaccurate, more rigorous testing of the pilot might compare individual performance on a careful reading test, e.g. the KET, with our test of expeditious reading, thereby establishing reading level as determined by CEFR as well as providing a point of contrast between ability in these two reading skills.

Theoretical implications and future research

Our pilot provides some empirical evidence that low proficiency language learners may be able to engage in the less cognitively demanding expeditious reading skills (skimming and scanning). This finding adds support to Weir and Khalifa (2008, p.9) model of reading types and their differing cognitive demands, which posits that these expeditious reading skills are cognitively less demanding than most other types of reading. On the other hand, our finding of no difference in scores on skimming and scanning items may suggest that expeditious reading may be unidimensional, or at least the different components may be of similar cognitive difficulty. However, without

knowing the time spent on each item focus, this is only a possible springboard to further research.

Our preliminary findings also highlight the potential under-representation of expeditious reading in the CEFR itself at the A2 level, an echo of perhaps a broader over-generalization and lack of specificity in the CEFR levels regarding speed and strategy as a reflection or reading purpose (Weir, 2005); this fact may mirror the CEFR's sociolinguistic, rather than psycholinguistic orientation (Weir, 2005; Alderson, 2007).

The absence of expeditious reading in A2 descriptors may have had knock-on effects in curricula and test design, with washback effects de-emphasizing these skills in the classroom and leading to potential subsequent difficulties since B1 and B2 learners are expected to engage in a new set of skills in which they have had much less practice as compared to careful reading.

7. Conclusions

We set out on this project to investigate not only ways of testing expeditious reading in lower level learners but to determine the extent to which this skill —absent from the CEFR framework at A2— may be within the skill set of this group. According to these preliminary results, it seems clear that this set of reading skills, restricted for our purposes to scanning and skimming, is within the ability of lower-level learners. This shows support for the order of difficulty proposed by Weir and Khalifa (2008) as discussed in the literature review, in which scanning for information and skimming for gist are designated as cognitively less demanding than most types of careful reading; likewise, it indicates a possible under-representation of this skill in the CEFR A2 descriptors.

As a pilot, our test cannot yet speak conclusively to these issues, but it has provided a sketch of low proficiency learners' abilities in certain types of expeditious reading. Further development and testing of this pilot test would help delineate this ability and its appropriate assessment.

8. References

Alderson, C. 2000. Assessing reading. Cambridge: Cambridge University Press.

Alderson, C. 2007. The CEFR and the need for more research. The Modern Language Journal 91(4): 659-663.

Bax, S. 2013. The cognitive processing of candidates during reading tests: Evidence from eye-tracking. Language Testing 30(4): 441-465.

Bernard, M., R. Baker and M. Fernandez. 2002. Paging vs. scrolling: Looking for the best way to present search results. Usability News 4(1).

Brown, J. 1988. Understanding research in second language learning: A teacher's guide to statistics and research design. Cambridge: Cambridge University Press.

Carver, R. 1992. Reading rate: theory, research and practical implications. Journal of Reading 36(2): 84–95.

Cohen, A. D. and T. A. Upton. 2007. 'I want to go back to the text': Response strategies on the reading subtest of the New TOEFL[R]. Language Testing 24(2): 209-250.

Cullen, P., A. French and V. Jakerman. 2014. The official Cambridge guide to IELTS: Students book with answers with DVD-ROM. Stuttgart: Klett.

Educational Testing Service. 2012. The official guide to the TOEFL test. McGraw-Hill Education.

Enright, M. K., W. Grabe, K. Koda, P. Mosenthal, P. Mulcahy-Ernt y M. Schedl. 2000. Toefl 2000. Reading framework: A working paper. Princeton, New Jersey: Educational Testing Service.

Grabe, W. and F. L. Stoller. 2002. Teaching and researching reading. Harlow: Pearson.

Grellet, F. 1981. Developing reading skill. Cambridge: Cambridge University Press.

Haladyna, T. and M. Rodriguez. 2013. Developing and validating test items. New York: Routledge.

Khalifa, H. and C. J. Weir., 2009: Examining reading: Research and practice in assessing second language reading. Cambridge: Cambridge University Press.

Koda, K. 2005. Insights into second language reading. New York: Cambridge University Press.

Ministerio de Educacion de Chile. 1999. Decreto Nº 112/99 (1º y 2º medio). Disponible en http://www.mineduc.cl/biblio/documento/dex112_99.pdf, [Consulta: 10/10/2015]

Moore, T., J. Morton and S. Price. 2007. Construct validity in the IELTS Academic reading test: A comparison of reading requirements in IELTS test items and in university study. British Council/IDP Australia IELTS Research Reports 11: 120-211.

Paran, A. (1991). Reading comprehension. Limassol, Cyprus: Burlington.

Robb, T.y B. N. Susser. 1989. Extensive reading vs. skills building in an EFL context. Reading in a Foreign Language 5: 239–251.

Rose, D. 2006. Report on observations of FCE and CAE time trials. Cambridge: Cambridge ESOL internal report.

Sanchez, C. A. and J. Wiley. 2009. To scroll or not to scroll: Scrolling, working memory capacity, and comprehending complex texts. Human Factors: The Journal of the Human Factors and Ergonomics Society 51(5): 730-738.

Schmidt, R. 2001. Attention. En P. Robinson (Ed.), Cognition and second language instruction. Pp.3-32. Cambridge: Cambridge University Press.

Shiotsu, T. 2010. Components of L2 reading: linguistic and processing factors in the reading test performances of Japanese EFL learners. Cambridge: Cambridge University Press.

Urquhart, A. H. y C. J. Weir, 1998. Reading in a second language: Process, product and practice. London: Longman.

Weir, C.J. 1983. Identifying the language needs of overseas students in tertiary education in the United Kingdom. PhD Thesis: University of London, Institute of Education.

Weir, C. J. 1993. Understanding and developing language tests. New York: Phoenix.

WEIR, C. J. 2005: Limitations of the common European framework for developing comparable Examinations And Tests. Language Testing 22(3): 281-300.

WEIR, C. J., Y. HUIZHONG AND J. YAN. 2000. An empirical investigation of the componentiality of L2 reading in English for academic purposes (Vol. 12). Cambridge University Press.

WEIR, C. J. AND H. KHALIFA. 2008. A cognitive processing approach towards defining reading comprehension. Cambridge ESOL: Research Notes 31: 2-10.

WEIR C. J., C. HAWKEY, R. GREEN AND S. DEVI. 2009a. The cognitive processes underlying the academic reading construct as measured by IELTS. British Council/IDP Australia IELTS Research Reports. 9: 157–189.

WEIR C. J., R. HAWKEY, A. GREEN, A. UNALDI AND S. DEVI. 2009b. The relationship between the academic reading construct as measured by IELTS and the reading experiences of students in their first year of study at a British university. British Council/IDP Australia IELTS Research Reports, 9: 97–156.

WEIR C. J., I. VIDAKOVIC AND E. DIMITROVA-GALACZI. 2013. Measured constructs: a history of Cambridge English language examinations 1913-2012. Cambridge: Cambridge University Press.

APPENDIX A

A model of cognitive processing of reading (Weir & Khalifa, 2008)

Creating an intertextual representation:
Construct an organised representation across texts

Text structure knowledge:
Genre
Rhetorical tasks

Creating a text level representation:
Construct an organised representation of a single text

Remediation where necessary

Building a mental model
Integrating new information
Enriching the proposition

General knowledge of the world
Topic knowledge
Meaning representation of text(s) so far

Monitor: goal checking

Inferencing

Goal setter
Selecting appropriate type of reading:

Careful reading
Local:
Understand sentence

Global:
Comprehend main idea(s)
Comprehend overall text
Comprehend overall texts

Expeditious reading
Local:
Scan/search for specifics

Global:
Skim for gist
Search for main ideas
and important detail

Establishing propositional meaning
at clause and sentence level

Syntactic Parsing

Syntactic knowledge

Lexical access

Lexicon Lemma:
Meaning
Word class

Word recognition

Lexicon Form:
Orthography
Phonology
Morphology

Visual input

APPENDIX B

Task features of relevant reading papers

Sources of Reading papers	Cambridge ESOL	
	KET (A2)	PET(B1)
Text length	250 (maximum)/text; 740-800 words/4 texts	550 (maximum)/text; 1450-1600 words/5 texts
Time constraints	35 items / 40 min	35 items/50 min
Types of reading & Response methods	Careful reading local: (multiple choice, multiple matching, multiple-choice cloze)	Careful reading local: (multiple choice, multiple-choice cloze); Careful reading global: (multiple choice); Expeditious reading local: (true/false); Expeditious reading global: (multiple matching)
Text purpose	Mainly referential	Mainly referential, emotive
Rhetorical task	Descriptive,narrative, instructive	Descriptive, narrative, expository, instructive
Genre type	Public signs & notices; newspapers & magazines; informational sources	Public signs & notices; newspapers & magazines; informational sources; personal messages
Grammatical resources	Flesch reading ease: 78.3 Flesch-Kincaid grade level: 5.5	Flesch reading ease: 64.7 Flesch-Kincaid grade level: 7.9
Lexical resources	Type-token ratio: .37	Type-token ratio: .30

(Source: Khalifa & Weir, 2009)

APPENDIX C

Item analysis statistics

Time allocation prototyping

Item	Item Facility	Item Discrimination
1	0.52	0.63
2	0.64	0.14
3	0.32	0.80
4	0.32	0.71
5	0.44	0.55
6	0.16	0.32
7	0.20	0.62
8	0.40	0.37
9	0.40	0.63
10	0.32	0.64
11	0.40	0.74
12	0.32	0.75
13	0.44	0.12
14	0.32	0.23
15	0.12	0.15
16	0.12	0.41
17	0.04	0.24

First timed pilot

Item	Item Facility	Item Discrimination
1	0.76	0.28
2	0.82	0.06
3	0.42	0.33
4	0.67	0.55
5	0.55	0.72
6	0.61	0.54
7	0.55	0.51
8	0.30	0.55
9	0.85	0.40
10	0.33	0.83
11	0.36	0.37
12	0.39	0.70
13	0.33	0.49
14	0.06	0.11
15	0.09	0.57

Second timed pilot

Item	Item Facility	Item Discrimination
1	0.33	0.61
2	0.71	0.46
3	0.58	0.45
4	0.63	0.53
5	0.71	0.59
6	0.75	0.59
7	0.46	0.48
8	0.29	0.74
9	1.00	n/a
10	0.38	0.62
11	0.54	0.21
12	0.50	0.45
13	0.50	0.47
14	0.54	0.33
15	0.17	0.75

APPENDIX D

Test Specifications

Rationale:

This test has been designed to help teachers assess students' expeditious reading skills. These skills, skimming and scanning, have been identified as an area of difficulty for L2 learners of English, and show a wider gap than careful reading skills when comparing native and non-native readers. Because of this, students might benefit from an earlier start in practice, particularly if they will be working or studying in an English speaking context or will need to take standardized English proficiency tests, which test these skills, in the future. It will help teachers to be able to identify whether or not their students are lagging behind in these skills.

Target test taker:

This test is aimed at A2 level students and is meant to be done in class as an exercise to help inform teachers how their cohort performs at expeditious reading operations.

Test use:

This test can be used as an in-class exercise to help teachers determine the degree to which they might need to focus on expeditious reading skills for a particular group.

Test format:

The test is done on a computer. The test involves reading and answering questions on three texts under a time limit.

Time Limit:

The test is to be done within 20 minutes. The timing starts after the instructions have been reviewed by the test taker. Once the time limit is reached, the test automatically ends and submits scores.

Instructions:

Because the test is focusing on a particular set of reading skills with which students might not be familiar, clear instructions in L1 are required to orientate test-takers to the task. Instructions outline the purpose of the test and that it should be done as quickly as possible. There is also an example text with two questions (gist - MC; scanning - SAQ) to help test takers get acquainted with the format before they begin.

exts:

Texts are between 450-500 words and have one of the following formats:

• Personal correspondence

• Public information

    – Brochure

    – Article

    – Advertisement

Texts have one of the following communicative purposes:

• To give information

• To give advice

Texts have a clear structure and are divided into paragraphs or sections.

Text difficulty is around B1 level, though a higher level of language is permitted if it is not a tested part of the text. As a general guideline, texts have the following range in the following measures:

- A Flesch reading score of above 70.

- A type token ratio of below .5

Item Types:

Each text is accompanied by five questions, positioned ahead of the text so that test-takers can identify what information they should search for when reading.

For each text:

- One question will focus on gist or skimming

  – This question will always be a Multiple Choice item type with one key and three distractors.

  – This question will often ask about the purpose of the text or the writer's / subject's attitude or feelings.

- Four questions will focus on scanning and search reading.

  – These questions can take the form of either

    Multiple Choice Questions (as above)
    - More often used with scanning questions

    Short Answer Questions
    - Best used with scanning questions

    - Sometimes presented as gapped sentences.

Scoring Method

Each question is weighted equally, for a total of 15 points. A key is provided to scorers on the answers for the multiple-choice questions and short answer questions. For the latter, a number of possibilities are delineated, and scorers are instructed to ignore spelling and grammatical errors.
Each test result is scored twice by two different scorers to ensure that all answers have been accurately rated and there is agreement on short answer question answers. There is no pass or fail mark for the test.

Feedback

Once tests have been scored, test results will be reported to the teachers of the groups with guidelines on how to interpret the results.