Title of Project: The Effects of Task Complexity on the Difficulty of L2 Reading Test: An Eye-tracking Study

Researcher: Jookyoung Jung
University College London
Jookyoung.jung@gmail.com

Research Supervisor: Dr. Andrea Révész
University College London
A.Revesz@ioe.ac.uk

Project Summary

The aim of the present research is to explore the relationship between task-related variations and difficulty of L2 reading assessment. Although a large number of factors have been investigated that affect the difficulty of L2 reading tests, including textual features of the reading passage (e.g., readability, topical structure, and length), test-takers’ topic familiarity with the reading passage, the level of required understanding (e.g., literal vs. inferential or local vs. global), and test format (e.g., open-ended, cloze, and multiple-choice), task complexity has not been considered as a potential factor contributing to test difficulty. How the online cognitive processes in which test-takers engage may relate to the quality of the reading comprehension outcomes has also been unattended. To fill these gaps in the literature, this project aims to examine test-takers’ cognitive processes while performing L2 reading assessment tasks, and the ways in which the cognitive processes may relate to the quality of the texts comprehension, which will further be used to establish the cognitive validity (Bax, 2013; Brunfaut & McCray, 2015; Khalifa & Weir, 2009) of the test.

Thirty-eight native Korean speakers participated in the study. They all completed two versions of reading assessment tasks. Following a 2x2 repeated-measures design, the test-takers were exposed to the two texts under a simple or complex task condition. Text order was counterbalanced across the test-takers. Both texts were divided into five segments. Each segment was presented on one page, aligned with the original format of the TOEFL tests. The reading tasks involved ordering parts of the segments and answering multiple-choice comprehension questions. The task complexity operationalization entailed manipulating the cognitive demands posed by the text-ordering task. Under the simple condition, each text segment was split into two subparts (A and B), whereas, under the complex condition, the segments were divided into four (A, B, C, and D). The test-takers were asked to determine the correct order of the parts under both the simple and complex conditions before answering the multiple-choice comprehension items. It was assumed that the task version, which required the re-ordering of more subparts would generate greater cognitive demands, given that reading comprehension is facilitated by the clarity and coherence of text structure (Meyer, 1975; Meyer & Freedle, 1984; Meyer & Ray, 2011). While performing the assessment tasks, the test-takers’ eye-movements were recorded with a mobile Tobii X2-30 eye-tracking system with a temporal resolution of 30 Hz. After completing the two reading sessions, eleven students were further invited to take part in stimulated recall protocols.
prompted by recordings of their eye-movements made during reading. They were instructed to verbalize what they were thinking while engaged in the original reading task.

Eye-tracking data were analysed with Tobii Studio 3.0.9 (Tobii Technology, n.d.). For each page, areas of interest (AOI) were defined for (a) the text and (b) the text and response options combined. Eye-movements captured on the text AOIs were used to extract indices associated with text reading processes, whereas AOIs for the text and response options combined were utilized as the basis for calculating measures of global processes during task performance. Then, drawing on Brunfaut and McCray’s (2015) work, in total, eye-movement indices of text and global processing were calculated based on eye-movement data such as eye fixations, during which the eye dwells on part of a text, and saccades, which occurs when the eye moves from one location to the next, obtained from Tobii Studio using R-script (McCray, 2016). The following indices were extracted from the eye-movement data: number of fixations, sum of fixation durations, median fixation duration, number of forward saccades (eye-movements from point x to point y where point y lies to the left of point x), median length of forward saccades, number of regressions (eye-movements from point x to point y where point y lies to the right of point x), median length of regressions, and proportion of regressions (The number of regressions divided by the sum of the number of both forward saccades and regressions).

The stimulated recall sessions were transcribed using the video-transcription software F5, Version 2.2. The transcripts were uploaded to NVivo 10.0.3 software for qualitative analysis. The researcher reviewed the transcripts and identified emergent categories in a bottom-up manner by annotating the data. After coding all the transcripts, a randomly selected subset of the video-recordings (13.6%) was watched and coded by a second coder, an expert in Applied Linguistics, in order to verify the reliability of the coding. Agreement between the researcher and the second coder was 90 per cent with a kappa of .71 (SE = 1.02, 95% CI [-.98, 3.06]), which was acceptable. Next, comments were further categorized depending on whether they concerned the simple or complex condition, and frequency counts were calculated for each code under each condition.

The results from mixed-effects modeling on the eye-movement indices revealed that the test-takers processed the texts more thoroughly under the complex than the simple condition. That is, the participants tended to fixate more on the assessment tasks when performing the complex versions. In addition, they fixated more frequently and for longer on the texts under the complex condition, as manifested in the significantly larger number of fixations and longer fixation durations for the texts. The numbers of forward saccades and regressive eye-movements further indicated that the participants engaged in more attentive and recursive processing of the texts. Successful completion of the complex tasks may have required closer inspection of the texts in order to arrive at an accurate understanding of each sentence, as well as the logical relationships among them. Consequently, the participants might have had to read the texts more carefully and thoroughly when completing the complex tasks, which was confirmed by the eye-movement data.

The analysis of stimulated recalls provided results compatible with eye-movements. On a global level, the test-takers reported that they perceived the complex task as more demanding. In particular, they more often recalled wrestling to order the segments and being unconfident about task completion. The significantly greater number of fixations during the task may represent the test-takers’ deliberate endeavours to process the text for accurate understanding, which was crucial to order the text segments coherently under the complex condition. The test-takers’ comments also revealed that, under the complex condition, they more frequently employed various reading strategies, such as skimming, careful reading and
searching for hints. They also recalled more extensive use of lexical cues, including keywords, signal words, pronouns and words that were mentioned for a second time. That is to say, they appeared to process the texts more intensively using diverse metacognitive strategies under the complex condition, which seems consistent with the longer duration of fixations, as well as the increased numbers of fixations, forward saccades, and regressions captured in the texts.

The task-complexity manipulation, however, had no significant influence on the reading comprehension scores. In other words, the effects of task manipulation appeared to be more readily observable in the test-takers’ cognitive processes, which might not necessarily surface in the reading comprehension scores. In this regard, the results from this study corroborate the need to investigate test-takers’ reading processes during performing assessment tasks in order to achieve a fuller understanding of the difficulty, as well as cognitive validity of a L2 reading test. More importantly, this study demonstrates that task complexity may operate as a potential factor contributing to the difficulty of a reading assessment, which calls for more empirical evidence. In a similar vein, a theoretical framework or an applicable taxonomy for analyzing the cognitive demands of a reading assessment item may need to be constructed based on accumulated research findings on the relationships between task complexity and test-takers’ reading processes.
References


Bates, D.M., Maechler, M., & Bolker, B. (2012). *lme4: Linear mixed-effects models using S4 classes*. R package version 0.999999-0.


177 Webster St., # 220, Monterey, CA 93940 USA
Web: www.tirfonline.org / Email: info@tirfonline.org


