

TEST-TAKERS' COGNITIVE PROCESSES DURING INTEGRATED WRITING WHICH USE MULTIPLE TEXTS AND GRAPHS AS PROMPTS -Preliminary Findings-

Mikako Nishikawa

Ph.D. Candidate,
Graduate School of Education
University of Bristol

This presentation is based on preliminary data for a doctoral thesis.

LTF Conference, 2016

Please do not cite without permission.

[Mail] mikako.nishikawa@bristol.ac.uk

Outlines

- Background
- Purpose
- Research Questions
- Methodology / Design & Instruments
- Methodology / Measurement
- Preliminary Findings (Data Analysis)
- Summary of Findings (for RQs)
- Limitations / Future Studies
- References / Acknowledgement



University of
BRISTOL

Background

03 December 2016

(1) What is integrated writing?

Integrated tasks which require multiple language skills are employed in a number of international exams.
(e.g. iBT TOEFL, IELTS)

Independent task Knowledge-Telling	Integrated task Knowledge Transferring
The essay answers the topic question based upon own experience or recollections from one's memory	The essay reports the content, facts, and ideas of the source texts by synthesizing the information.

(Scardamalia & Bereiter ,1987)

3



University of
BRISTOL

Background

(2) Why do we test integrated tasks for writing?

- Integrated writing is a fundamental skill in academic writing which is even **difficult for L1 learners** (e.g. Cumming, Rebuffot & Ledwell, 1989; Shi, 2004; Hyland 2005; Delaney 2008).
- Integrated writing tasks give a **better prediction of** whether a learner will be **a novice or an advanced writer in a real-life academic setting** (Plakans, 2008).



University of
BRISTOL

Background

(3) Previous research on integrated writing

Some of the potential elements that may affect integrated writing include:

- The **use of the first language and target language** for summarization (e.g. Yu, 2008)
- Familiarity of **the discourse types of a source text** (e.g. Delaney, 2008; Yu, 2009) and **graph types** (e.g. Yu & Lin, 2014)
- The effects of **source text borrowing** (e.g. Cumming, et al, 2005; Weigle & Parker, 2012)
- Yu, Rea-Dickins & Kiely (2015) investigated the cognitive processes involved for summative writing with graphic information in the case of **IELTS Writing Task 1**.

5



University of
BRISTOL

Background

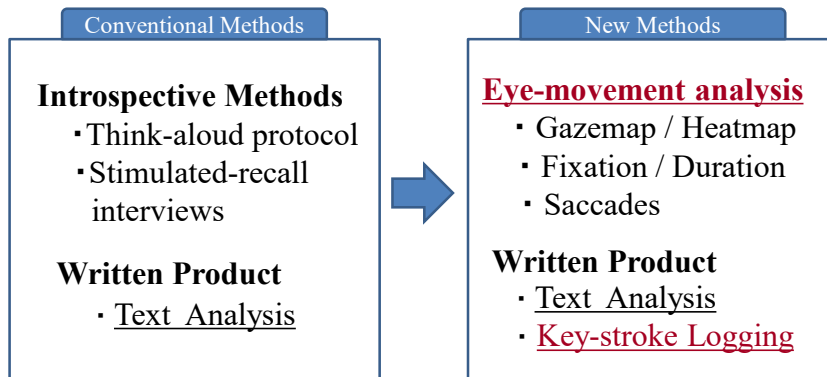
03 December 2016

(4) Research Gaps

- Previous studies suggest complexity of the reading into writing construct **is quite different from that of reading and writing as separate skills** (Delabey, 2008; Bachman, 2002).
- Little research has been conducted on **the use of graphic information** for integrated writing. (e.g. Yu, 2009; Yu, Rea-Dickins & Kiely, 2012; Yang, 2012, Yu & Lin 2014).
- **Empirical research on the cognitive processes** of integrated writing in **L2 context** are still at scarce (Révész, 2014).

6

(5) Research methods on integrated writing



7

(6) Latest research on language assessment using eye-tracking and keystroke analysis

- Bax (2013) investigated the cognitive elements of the **IELTS Reading** test using an eye-tracker to investigate readers' higher order, post-lexical processing behavior.
- Révész & Lee (2015) studied underlying L2 pausing and revisions behaviors using **IELTS Writing Task 2**. (←Independent writing task)
- Brunfaut & McCray (2015) used eye-tracking to investigate text processing according to test-takers' characteristics (language proficiency) in the case of **Aptis Reading**.



8

Purpose

- To gain insights into the cognitive processes of L2 writers when synthesizing information from source texts and graphics to produce a written product (essay).

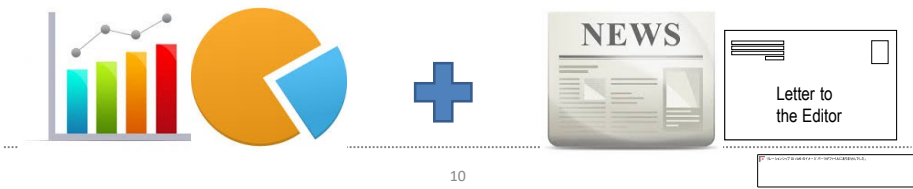
Eye-tracking as instrument

- Data from an eye-tracking device will help investigate cognitive processes of integrated writing in more depth.



TASK B (40-50 minutes long)

Your teacher has asked you to write an essay for class using the information below. Describe the situation concerning schools in Greenhill and summarize the main points about the solutions that have been suggested. In your conclusion, say which of the solutions you think would work the best based on the reasons given. You should write about 200 words.



Research Questions

RQ: What are the key variables that affect cognitive process of integrated writing?

- 1. To what extent do test takers incorporate information from multiple texts?
- 2. To what extent do test takers incorporate information from graphs?
- 3. To what extent do features of graphs affect the cognitive processes of writing given texts and graphic information as prompts?
- 4. What role does language proficiency play in integrated writing?
- 5. What kinds of test-taking strategies are used for integrated writing?



Methodology (Research Design)

A Mixed-Methods Approach

- Explanatory Research Method
 - QUAN. data will help assess the trends and relationships
 - QUALI. data will help explain the mechanism or reasons behind the results and trends

Participants

- 42 Japanese high school students
(Male=10, Female=32)

Grade 10	n=21
Grade 11	n=9
Grade 12	n=12

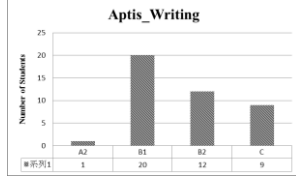
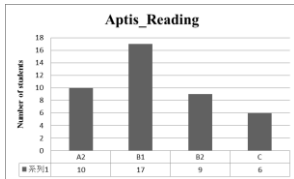
Data Collection (Procedure)

- Phase I: Aptis Reading, Writing, and Grammar & Vocabulary tests (n=54)
- Phase II: Two TEAP Writing Tests for eye-tracking (n=42)
- Phase III: Surveys for test-taking strategies (n=41)
- Phase IV: Focus group discussions (n=24)





Participants' Language Proficiency (Reading/Writing)



		APTIS_Score_Reading	APTIS_Score_Writing	Grammar & Vocabulary
APTIS_Score_R	Pearson Correlation	1	.574**	.703**
	Sig. (2-tailed)		.000	.000
	N	42	42	42
APTIS_Score_W	Pearson Correlation	.574**	1	.791**
	Sig. (2-tailed)	.000		.000
	N	42	42	42
Grammar_Vocabulary	Pearson Correlation	.703**	.791**	1
	Sig. (2-tailed)	.000	.000	
	N	42	42	42

** Correlation is significant at the 0.01 level (2-tailed).



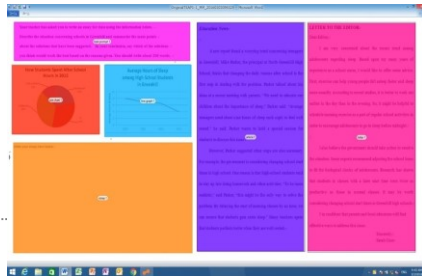
Data Analysis

Overall Trends for 40 minutes Recordings 1: Graphic Information

Analysis:

- Recordings which had little gaze percentage (5% less) were omitted
- Paired-Sample T-test for 40 minutes of recordings (N=38)

Areas of Interests for Test 1



Areas of Interests for Test 2

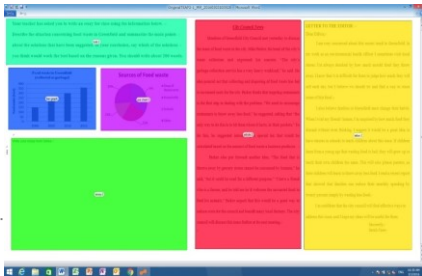


Table 1. Pair-Sample T Test Results Please see my handout ☺

Findings 1. Line Graph vs. Bar Graph

-Time to First Fixation for Line Graph and Bar Graph

It took longer for students to have their eyes first fixated on Line Graph (Test 1) than on Bar Graph (Test 2)

-Implication:

The results could be due to the locations of the graphs being placed in the two tests.

-Means for Visit Counts

Students had a greater Visit Count for Line Graph (Test 1) than they did for Bar Graph (Test 2).

-Implication:

The students studied a line graph more than they did a bar graph.

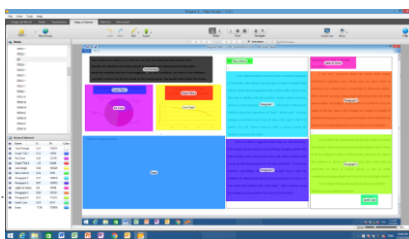
Data Analysis

Overall Trends for 40 minutes Recordings 2: Effects of Language Proficiency

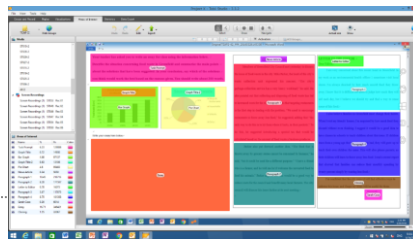
Analysis:

- Regression Analysis (Aptis Reading/Writing scores as predictors)
- Independent-T Test for Reading **【B1 below(N=24) vs. B2 above(N=14)】**
- Mann-Whitney U Test for Writing **【B1 below(N=19) vs. B2 above(N=19)】**

Sub-Divisions of Areas of Interests for Test 1



Sub-Divisions of Areas of Interests for Test 2





University of
BRISTOL

Data Analysis

Table 2. Regression Analysis Please see my handout ☺

Findings 2. Writing is a better predictor of the values in eye-tracking measurement

Caution:

- The results remain just as reference.
- Regression analysis is not an appropriate test for non-parametric data when the assumption of normality violated.



University of
BRISTOL

Data Analysis

Table 3. Significant Differences Between B1 below and B2 Please see my handout ☺

Findings 3. No variable found significant for reading in Test 1 except Ratio of Sum of Fixation Duration for Task Prompt

-Implication:

- B2 above students did not have to spend more time on the task prompt in Test 2
- B1 below students had taken more time reviewing the task prompt in Test 2
 - = Resulting in the statistically significant difference between the two groups

Findings 4. Among the AOIs, the most significant differences were found in the area of the Essay.

- Means of Fixation Duration, Fixation Counts for Essay for Reading and for Writing
- B2 above in reading and writing tended to spend more time in composing essay





University of
BRISTOL

Writing Processes

This Presentation
will focus on here.

(1) Preparing to Write: (Measurement)

Total Fixation /Total Duration in AOIs, Fixation Counts in AOIs (First/Last 5min),
AOI switches (First 10min), Survey Results,

- Task representation Did you understand what is required for the task?
- Macro-planning How did you plan when you started writing essay?
- Micro-planning How did you decide which information to include?

(2) Translation and Writing: (Measurement)

Total Fixation Duration in AOIs, Average Forward Saccades, Average Regressions, Time to
First Keystroke Logging, Number of Words, Number of Pauses

- Describing What kind of strategies did you use during the test?
- Synthesizing What did you often do while you were writing?
- Choosing positions
- Giving reasons

(3) Monitoring and Revising: (Measurement)

Total Fixation Duration in AOIs, Fixation Counts in AOIs (Last 10 min), Survey Results,
Number of Revisions (Backspace), Number of Key Movement (Right/Left/Up/Down)

- Task requirement What did you do after you finished writing the essay?

(Shaw & Weir, 2007)



University of
BRISTOL

Data Analysis

Findings 5: Task Representation

Did you understand what is required for the task?

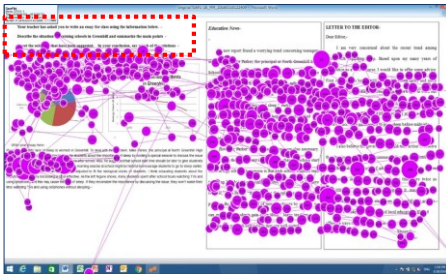
One-way ANOVA test (Writing)

A posthoc test (Dunnnett T3) revealed which groups had statistically significant differences since the homogeneity of variance ($p < .05$) is not assumed.

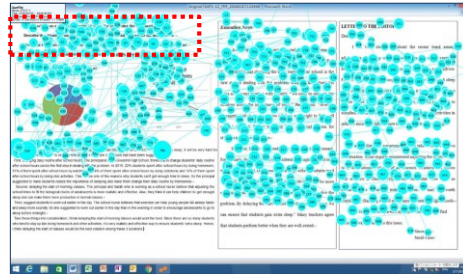
- Q3_1 Did you understand the instruction on how to write your essay?
F(2, 36)=9.028, $p=.001$
B1 level (M=.63, SD=.496) and B2 level (M=.00, SD=.000)
- Q4_1 Did you identify the purpose of the essay?
F(2,36)=6.034, $p=.005$
B1 level (M=.63, SD=.597) and B2 level (M=.09, SD=.302)
B2 level (M=.09, SD=.302) and C1 level (M=.1, SD=.333)

Findings (6): First 5 minutes of Recording (TEST 1)

(ST030)
Reading CEFR level C
Writing CEFR level B2



(ST037)
Reading CEFR level B1
Writing CEFR level C



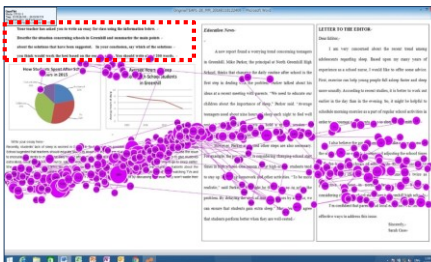
- ST030 went straight into the graphs and source texts.
- ST037 began by reading the task prompt.

21

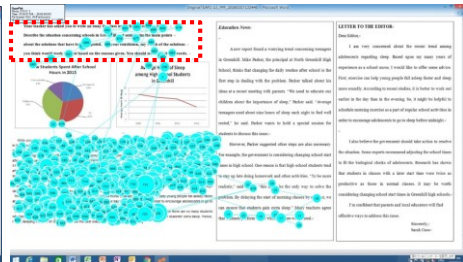
03 December 2016

Last 5 minutes of Recording (TEST 1)

(ST030)
Reading CEFR level C
Writing CEFR level B2



(ST037)
Reading CEFR level B1
Writing CEFR level C



- ST030 kept referring to source texts and graph titles while writing.
- ST037 went back to the task prompt to double check the task requirements.

22

Findings (7): Macro-planning

How did you plan when you started writing essay?

Q7_1 Did you make an outline BEFORE writing your essay?

Yes. (ST037) No. (ST030)

Q7_2 Did you decide how many paragraphs there should be in your essay?

Yes. (ST037) No. (ST030)

(ST030)

Reading CEFR level C

Writing CEFR level B2

- More simple mechanics errors (misspells) found for ST030
- Total number of source use is greater for ST037
- The words lengths are also greater for ST037
- Synthetic complexity is also greater for ST037

“Recently, students' lack of sleep is worried in Greenhill. To deal with this problem, Mike Parker, the principal at North Greenhill High School **sugested** that teachers should **educate students about the importance of sleep** by holding **a special session to discuss the issue** to encourage students to **change the daily routine after school**. Also, he suggested that school start time should be later to give **studends** extra sleep. Sarah Case, a school nurce, said that morning **exacise** at school might be helpful to encourage students to go to sleep earlier. She also believed that school times should be adjusted to **fit the biological clocks of students**. I think educating students about the importance of sleep by discussing is most effective. As the left **fegure** shows, many students spent after school hours **watching TVs** and using **cellphones** and this may cause the lack of sleep. If they reconsider the importance by discussing the issue, they won't waste their time **wathing** TVs and using cellphones without sleeping.”

(ST037)

Reading CEFR level B1

Writing CEFR level C

“In the school in Greenhill, the average hours of sleep is gradually decreasing since 2000. Without enough sleep, it will be very hard for students to perform better. To improve this situation there are 3 solutions that have been suggested.

First, **changing daily routine after school hours**. The principal at North Greenhill High School, thinks that to change students' daily routine after school hours can be **the first step in dealing with the problem**. In 2015, 22% students spent after school hours by doing homework, 31% of them spent after school hours by watching TV, 34% of them spent after school hours by using cellphone and 13% of them spent after school hours by doing club activities. This can be one of the reasons why students can't get enough time to sleep. So the principal suggested to make students realize the importance of sleeping and make them change their daily routine by themselves.

Second, **delaying the start of morning classes**. The principal and Sarah who is working as a school nurse believe that **adjusting the school times to fit the biological clocks of adolescents** is more realistic and effective. Also, they think it can help children to get enough sleep and can make them more productive in normal classes.

Third, suggest students to work out earlier in the day. The school nurse believes that **exercise can help young people fall asleep faster and sleep more soundly**. So she suggested to work out earlier in the day than in the evening in order to encourage adolescents to go to sleep before midnight.

Take these things into consideration, i think delaying the start of morning classes would work the best. Since there are so many students who tend to stay up late doing homework and other **activities**, it is very realistic and effective way to ensure students' extra sleep. Hence, I think delaying the start of classes would be the best solution among these 3 solutions.”



University of
BRISTOL

Data Analysis

03 December 2016

Findings (8) Micro-planning

How did you decide which information to include?

AOI Switches for 10 minutes of recording (Test 1)

ST030-1	P1	P2	P3	P4	PX	# of AOI Switches
P1	0	5	6	1	28	40
P2	3	0	1	0	7	11
P3	5	0	0	0	8	13
P4	0	1	0	0	6	7
PX	32	5	5	6	0	48

ST030

Reading CEFR level C

Writing CEFR level B2

There is an indication that ST030 did not read P4 after P3

ST037-1	P1	P2	P3	P4	PX	# of AOI Switches
P1	0	2	1	0	32	35
P2	0	0	1	5	5	11
P3	2	0	0	1	5	8
P4	0	4	1	0	2	7
PX	33	5	5	1	0	44

Both students viewed P1 the most.

ST037

Reading CEFR level B1

Writing CEFR level C



Finding (9): Micro-planning

How did you decide which information to include?

Reading Behaviors: Saccades/Regressions/Return Sweeps while reading P1

Test 1, Paragraph 1	Saccades	Regressions	Return Sweeps
ST030 (C/B2)	239 letters	121 letters	17 letters
ST037 (B1/C)	202 letters	119 letters	36 letters

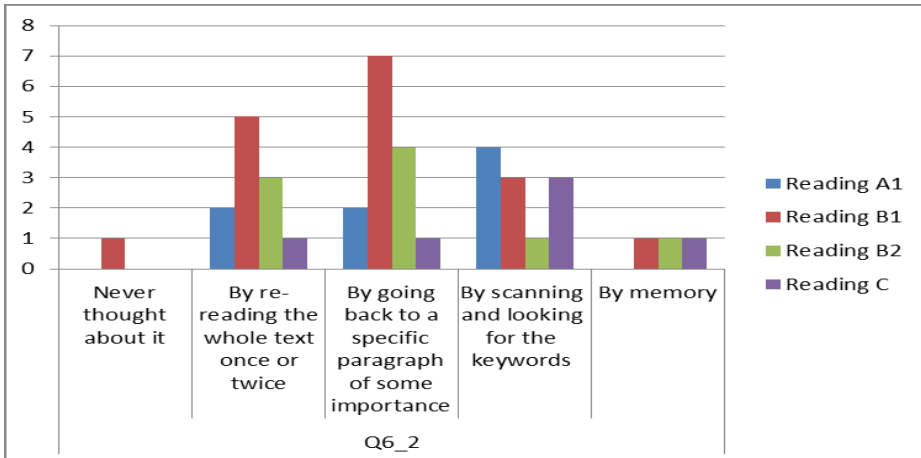
- The greater saccades for ST030 confirms she is a better reader than ST037

27



Finding (10): Micro-planning

Q6_2: How did you decide which information to include from the texts?



28



University of
BRISTOL

Summary of Findings (preparing to write) Task representation/Macro-planning/Micro-planning

RQ1 Effects of Source Texts

- Less successful participants depended on the first paragraph more than other paragraphs
- More successful participants read paragraphs in order (at least once) and scanned the keywords to find the information in texts

RQ2. Effects of Graph Information

- Most students found it easy to interpret graph information, but difficult to summarize the trends in English. Consequently, the majority of them did not plan to include information from graphs as much as from texts at the planning stage

RQ3. Effects of Features of Graph Types

- Students had a greater Visit Count for Line Graph (Test 1) than they did for Bar Graph (Test 2)
- Most students found the Pie Chart easier than the Line Graph (Test 1) and Bar Graph (Test 2)

RQ4. Roles of Language Proficiency in Integrated Writing

- First 10 minutes were very crucial for task representation and macro-planning based on evidence from cases studies of students of different language levels (B2 vs C)

RQ5. Test-taking Strategies for Integrated Writing

- Significant differences were found between B1 and B2 in
(1) Understanding the task requirement; (2) Identifying the purpose; and

29



University of
BRISTOL

03 December 2016

Limitation

- The current study only had a few A2 level students.
- Given small sample sizes, the findings cannot be used for generalization.
- Data from the eye-tracking device is overwhelmingly large and complex.
- Depending upon basis or applied research, the design needs to be more tightly controlled.

Future studies

- Individual cases will need to be examined in order to investigate what really happened during integrated writing processes.

30



References

- Please see my handout 😊

Acknowledgement

- ☐ I would like to express my deep gratitude to my supervisor, Dr. Guoxing Yu for his patient guidance and enthusiastic encouragement.
- ☐ I would like to also thank Dr. Kazanina and Dr. Isaacs for useful critiques of this research work.
- ☐ Many thanks are extended to the British Council, the Eiken Foundation of Japan and The International Research Foundation for Language Education (TIRF) for research funding, and
- ☐ Finally, I would like to thank to the teachers and their students for their participation in this study.



Thank you for listening!

Mikako Nishikawa

Ph.D. Candidate, Graduate School of Education, University of Bristol

This presentation is based on preliminary data for a doctoral thesis.

Please do not cite without permission.

[Mail] mikako.nishikawa@bristol.ac.uk



The International Research Foundation
for English Language Education

