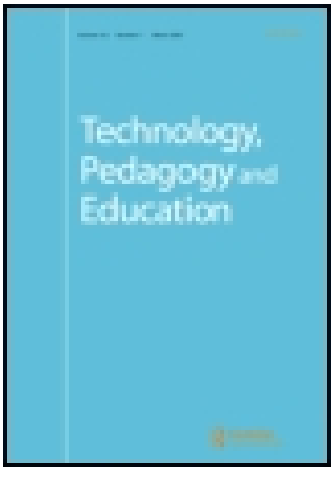

Technology, Pedagogy and Education
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/rtpe20

Effects of test media on different EFL test-takers in writing scores and in the cognitive writing process

Xiao-Ling Zou[a] & Yan-Min Chen[b]
[a] Research Academia for Linguistics, Cognition & Application, Chongqing University, Chongqing 401331, China
[b] School of Foreign Languages and Cultures, Chongqing University, Chongqing 401331, China
Published online: 21 Jan 2015.

Technology, Pedagogy and Education, 2015
http://dx.doi.org/10.1080/1475939X.2014.954140

Routledge
Taylor & Francis Group

Effects of test media on different EFL test-takers in writing scores and in the cognitive writing process

Xiao-Ling Zou[a]* and Yan-Min Chen[b]

[a]*Research Academia for Linguistics, Cognition & Application, Chongqing University, Chongqing 401331, China;* [b]*School of Foreign Languages and Cultures, Chongqing University, Chongqing 401331, China*

(*Received 23 July 2013; final version received 15 December 2013*)

Introduction

In recent years, there has been a growing interest in developing and using computer-based tests in many educational settings such as science, psychology and mathematics. Delivering assessments via computers in the domain of educational assessment is becoming increasingly prevalent along with the changes in assessment methodologies reflected in pedagogy (K. T. Chen, 2012; Chua & Don, 2013; Genc, 2012; Hsiao, Tu, & Chung, 2012; Organisation for Economic Co-operation and Development, 2010; Piaw, 2012). Computer-based testing, as a catalyst for change, has brought about a transformation in learning, pedagogy and curricula in educational institutions (Scheuermann & Pereira, 2008).

In language instruction, the interest in applying computer-based testing in educational assessment in schools and educational institutions has also been felt in recent years. More and more traditional paper-based tests such as TOEFL (Test of English as a Foreign Language) and GRE (Graduate Record Examination) have been redeveloped as computer-based tests. With more and more applications of computers in education, computer-based tests demonstrate more advantages over traditional paper-based tests in scoring, administration, score analysis (Li & Kong, 2009), feedback, score reports (Zou & Zhang, 2013), design and diagnosis (He & Tymms, 2005).

*Corresponding author. Email: xiaolingzou@cqu.edu.cn

To adjust to this tendency, China has been trying out the computer-based College English Test Band-4/6 (CET-4/6) in some universities and colleges since 2008, which is now being expanded to the entire nation. As one of the pilot universities of the CET-4/6 reform, the authors' university has not only participated in the pilot project of the computer-based CET-4/6, but also developed its own web-based College English testing system to evaluate students' language proficiency.

However, in the process of practice, some students reported that they preferred the traditional paper-based to the computer-based writing test, since they felt more nervous and anxious before the new test medium, which might result in poor writing performance.

Out of consideration for minimising the adverse effects brought about by the new test medium and for the purpose of promoting the development of the web-based language test, this research explores, from the perspective of EFL (English as a Foreign Language) learners, the effects of test media on the performance of test-takers. In particular, it examines how computer familiarity influences test-takers' writing scores and the cognitive writing process in computer-based writing tests. Recommendations to reduce the negative impact are proposed. It is expected that such suggestions can help EFL learners develop special strategies for web-based writing tests and provide some practical values for the spread of computer-based writing tests in EFL learning.

Literature review

Research on computer-based writing began as early as the 1980s, when the word processor was first in use. Interest was paid mainly to the influences of the word processor on students' composition, revising or editing behaviours on computers (Goldberg, Russell, & Cook, 2003; Imhof, Vollmeyer, & Beierlein, 2007; Leeson, 2006; Norris, Pauli, & Bray, 2007). These empirical studies were mainly focusing on the comparability and comparative study of computer-based and paper-and-pencil tests in writing scores and the writing process, as well as the effects of computer familiarity on writing.

Effects of test media on writing scores

Some studies (Russell & Haney, 1997; Russell & Plati, 2000) show that students obtained higher scores on computer-based writing tests, while other studies among the GMAT (Graduate Management Admissions Test) and the TOEFL test-takers (Breland, Lee, & Muraki, 2005; Bridgeman & Cooper, 1998; Wolfe & Manalo, 2004) revealed the opposite result, indicating that students obtained higher scores for texts composed by hand than for those composed on computers. As the above studies had not controlled the potential scoring biases, one researcher (H. K. Lee, 2004) word-processed the paper essays with their spelling and grammatical errors on to the computer and mixed them with the computer-written essays for scoring, with the trained scorers not being informed of the existence of the word-processed version. The result showed that there was no significant mean difference in the overall scores between the two test modes. However, other researchers (Powers, Fowles, Farnum, & Ramsey, 1994) not only converted the handwritten essays to word-processed format but also changed the original word-processed essays into handwritten format to examine whether there was a difference in scores on these

two test media. The result revealed that handwritten essays received higher scores than word-processed ones, regardless of the test modes in which the essays were originally produced. Still other researchers (J. Chen, White, McCloskey, Soroui, & Chun, 2011; Horkay, Bennett, Allen, Kaplan, & Yan, 2006; MacCann, Eastment, & Pickering, 2002; Organisation for Economic Co-operation and Development, 2010) reported no statistical difference in scores between the two writing modes, which supported the assumption that test media did not have any influence on test-takers' writing scores. These studies compared the test-takers' scores of the written products in the two test media, but the results are contradictory.

Effects of test media on the writing process

A number of studies have investigated the effects of different test media on the writing process, but no consensus has been reached. Some studies note that computers might either improve or interrupt the cognitive writing process (Bridwell, Sirc, & Brooke, 1985; MacArthur, 1988; Neuwirth, 1990; Owston, Murphy, & Wideman, 1992), but they mainly focused on first-language (L1) writers and the effects of word-processing on the writing process.

Y.-J. Lee's study (2002) of six Korean students suggested that there were differences in composing process at the level of individuals. Students spent more time planning on paper than on computer and the difference was statistically significant. In addition, more time was spent in producing text on the computer. Accordingly, he inferred that planning and text producing on the computer appeared to be more complex than on paper and the incorporation of computers into writing tests involved a new way of thinking about the composing process. This conclusion, however, was based on small samplings.

Analysis of candidates' responses to a theory-based questionnaire about internal processing (Weir, O'Sullivan, Jin, & Bax, 2007) indicated a similar pattern between the cognitive processes involved in writing on computer and writing with paper and pencil. The results of student satisfaction surveys (Whithaus, Harrison, & Midyette, 2008) reveal that the students preferred keyboarding to handwriting, because keyboarding was faster, more familiar, legible and convenient, errors were more easily spotted and the computerised version was cleaner since they could edit and proofread more effectively on screen. Also the comparability of paper and computer versions of a functional writing assessment administered to adults aged 16 and older was investigated (J. Chen et al., 2011), whose results show that adults performed better overall and on most aspects of the writing tasks when writing on paper rather than on a computer. Relatively little research has been centred on all stages of the cognitive writing process.

Effects of computer familiarity on writing

According to McDonald (2002), individual differences such as computer familiarity, computer anxiety and computer attitudes could potentially influence test-takers' performance. Therefore, the degree to which test-takers have experience of computer use has been discussed regarding computer-based tests, with the increasing use of computers in language testing. Vispoel, Rocklin, and Wang (1994), as well as Fulcher (1999), reported no significant effects of computer familiarity on language tests administered on computer and on paper, while Taylor, Jamison, Eignor, and

Kirsch (1998) found a small but significant effect of computer familiarity on computer-based TOEFL, with students who were less familiar with computers receiving lower scores. But through giving a computer tutorial to the students, Taylor, Kirsch, Eignor, and Jamieson reported that 'no evidence exists of an adverse relationship between computer familiarity and computer-based TOEFL test performance due to lack of prior computer experience' (1999, p. 220).

Another study (Russell, 1999) about the effects of keyboarding ability on the performance of eighth-grade children in tests of math, science, writing and comprehension found that the students with low levels of keyboard skills were disadvantaged by a computer-writing test compared with students with similar low levels of keyboard skills taking a paper test. The opposite effect was observed for students with high levels of keyboard skills, who performed better on computers than in the paper examinations. In a subsequent study, however, Russell and Plati (2000) argued that eighth- and tenth-grade students performed better on computer-based writing tests regardless of keyboarding speed. Horkay et al. (2006) countered these arguments by pointing out that the use of word processors in a writing test did not influence the overall scores of essays written by students with medium to high levels of computer familiarity, but negatively affected essay scores of students with a low level of computer familiarity.

It can be seen that the studies reviewed here provide contradictory results. Some studies support the view that students' computer familiarity or experience could negatively affect their performance on computer-based tests while other studies have found no effects. However, with the advances of computer technology, computers are now more 'user friendly' than they were 10 or 20 years ago. In addition, more and more people are now exposed to computers in all areas of life, and students are obtaining more computer experience now. Therefore, the continual changes in computer experience make it difficult to summarise the effects by using past studies, thus there is a need to carry out more research upon the effects of computer familiarity on different test-takers.

In China, the studies for computer-based tests are mainly related to reading (Li, 2006) and speaking (Gao, 2007; Sun, 2007), as well as the validity of the computer-based language test (Jin & Wu, 2010; Li, 2006; Li & Kong, 2009), while empirical studies in the field of computer-based writing tests are rarely dealt with.

To sum up, although some comparative studies have been conducted on computer- and paper-based writing tests from the aspects of writing scores, the writing process and the effects of computer familiarity on writing, most of them were compared from a macro-perspective and their conclusions were not consistent. Few comparisons between different groups with different computer familiarity were made from the microscopic view, especially of the cognitive writing process in different writing stages, and the reasons for the differences were less explored. It is therefore necessary to carry out further research upon the effects of test media on the performance of test-takers at all stages of the writing process and the main factors affecting test-takers' writing scores and cognitive writing process in computer-based writing tests. This study is aimed at achieving these by addressing the following research questions:

(1) Are there any differences in the test-takers' overall scores between computer-based and paper-based writing tests? If the answer is yes, are the differences of their levels in computer familiarity reflected in their writing test scores?

(2) Do the test-takers with different levels of computer familiarity perform differently in the cognitive writing process in the computer-based writing test? If the answer to the question is yes, what are the main factors affecting the test-takers' performance in the computer-based writing test?

Theoretical framework

In early research, writing was regarded as a linear process which was divided into three linear stages including pre-writing, writing and rewriting (Rohman, 1965). However, this model was criticised by researchers who suggested that composing was not a process proceeding with clear-cut stages, beginning with one activity and ending in another. Sommers (1980), after having studied experienced and inexperienced writers, challenged the linear stage model and redefined the revising as a recursive process which, instead of a separate and final process of writing, relates to other writing processes. Gradually, more and more researchers accept Sommers' view that the composition process is recursive.

Flower and Hayes (1981), on the basis of verbal protocol analysis, put forward the famous cognitive process theory of writing, emphasising the non-linearity in the writing process, which consisted of three major components such as task environment, writer's long-term memory and writing process. The writing process was divided into three basic stages of planning, translating and reviewing, where planning was further split into three sub-stages: goal-setting, generating ideas and organising ideas.

This cognitive writing process was a major focus for the understanding of writing in the 1980s. From then on, more researchers have tried to explore other factors that influence the writing process. Bereiter and Scardamalia (1987) pointed to two different processing models in terms of recursive and interactive stages. One was the knowledge-telling model with a relatively low level of cognitive complexity from oral language experiences into written forms, and the other was the knowledge-transformation model with a relatively high level of cognitive complexity, involving more complex processing tasks such as considering the relevance of information, arranging information and logically organising ideas. Kellogg (1996) also noted the non-linearity of the writing process and sorted it into three stages: formulation, execution and monitoring, which were aligned with the three basic stages of writing, with formulation corresponding to the stage of planning and translating ideas, execution to programming and executing written products, and monitoring to examining the match of ideas and products. Hayes (1996) extended Flower and Hayes' (1981) writing process model to two main elements: the task environment and the individual. The task environment was divided into the social and the physical environment, with the former being subdivided into the audience and collaborators, and the latter into the text written and the composing medium. Thus Hayes' new model illustrated that the composing medium (i.e. paper vs. computer) played an important role in the task environment.

Grabe (2001) argued that these cognitive processing models in writing did not sufficiently consider linguistic knowledge, and therefore might not completely account for all second-language (L2) writing processes. He regarded language proficiency as a critical element of L2 writing. Besides, many researchers have studied the writing process within a communicative framework. For example, Celce-Murcia

and Olshtain (2000) examined the factors of contexts and language knowledge that impacted the writing process. They depicted the writing process in terms of interaction, including bottom-up and top-down processes. The bottom-up process involved such language knowledge as grammar, spelling, punctuation, vocabulary and cohesion, while the top-down process was associated with writers' previous knowledge on content, awareness of goals and knowledge of discourse. This writing process model emphasises the interactive feature.

To sum up, there is general agreement that the writing process has three basic components: planning, translating and reviewing. In addition, various factors can influence each stage of the writing process. For the purpose of this study, the authors mainly took the views of Flower and Hayes (1981) and Hayes (1996) on the writing process as the theoretical framework of this research.

Experimental study

Participants

The participants of this study were 216 second-year students of non-English majors from the authors' university. Their English level had been evaluated as Band 3 (the third level of College EFL testing) in the placement test they took when they entered the university. All the participants had taken part in web-based language testing organised by the university at least twice during the past two years, so they were familiar with the testing system. In the experiment, all the participants were required to complete carefully two comparable writing tasks under two different conditions. One half of the test-takers finished the first writing task on paper and the second on the computer. On the other hand, the other half wrote the first composition on the computer and the second one on paper. For further data analysis, valid participants were then sorted on the basis of their computer familiarity and classified into three groups respectively, called Group 1 (students with low level of computer familiarity), Group 2 (students with medium level of computer familiarity) and Group 3 (students with high level of computer familiarity). Since the two writing tests were conducted respectively in two class hours as part of the course requirement, all of the 216 students took part in them, whereas the two questionnaires were completed on a voluntary basis. The students who did not want to participate were able to withdraw. Besides, there were invalid questionnaires after the survey, so the final valid samples were 208 students.

Instruments

Questionnaires

Before writing on the computer, all the test-takers were required to complete a computer familiarity questionnaire based on Taylor et al.'s measurement model (1998). For the purposes of this study, the authors selected 11 items from the 23 items to distinguish the level of computer familiarity, including computer experience, frequency of use, type of use, access to computers, attitudes towards computers and related technologies. Taking the computer's popularity in language teaching into consideration, the authors modified the four choices of some of the items by raising the requirement of computer familiarity to a higher level (see Appendix 1). For each question, a single score was used to mark the four responses ranging from 1 to 4.

Thus the range of possible scores was from 11 to 44. After the students finished the questionnaire, the authors summed the scores of all the responses and then classified them into three levels: low, medium and high. Those whose total scores were in the range of 11 to 22.4 were defined as possessing a low level of computer familiarity; those whose total scores were in the scope of 22.5 to 32.4 were considered as having a medium level of computer familiarity; and those whose total scores were over 32.5 were grouped into the category of high level of computer familiarity. Accordingly, the students were classified into three groups, with 63 students in the high level, 87 in the medium level and 58 in the low level.

The computer-based cognitive writing process questionnaire, delivered after the computer-based writing test, contains two parts apart from the test-takers' personal information (see Appendix 2). In the first part, the 34 items for recording the test-takers' cognitive writing process were selected from the questionnaire designed by Weir et al. (2007). According to the research purpose, the wording of some items was slightly adjusted and the word 'computer-based' as a prerequisite was added before 'cognitive writing processes'. The test-takers were required to give responses to each item by selecting such alternatives as 1 = strongly disagree, 2 = disagree, 3 = neither agree nor disagree, 4 = agree and 5 = strongly agree. The second part (item 35) was to sort the factors affecting the test-takers' performance in the computer-based writing test in order of effect degree from high to low based on their own experiences.

Writing tests

According to McCutchen (1986), prior knowledge could have significant impact on the quality of writing; therefore, the choosing of the writing tasks was based on the following principles: first, all the participants are familiar with the topics; second, the style of the selected topics is argumentation since it needs more complex processing than other types of writing such as narration and can distinguish the writers' performance clearly (Grabe & Kaplan, 1996); third, the writing tasks are similar in difficulties and methods. Thus the two topics 'Reading Selectively or Extensively?' and 'Should the University Campus Be Open to Tourists?' had been chosen, for they had the same reasoning – a choice between two situations, and a discussion of advantages and disadvantages of the choice.

Before the test, the participants were informed of the purpose of the research and the requirements. After that, the participants began to take the exam. The time limit was 40 minutes and the word requirement was no less than 200 words for each of the writing tasks. The CET-4 holistic scoring guide was used as the rating scale for the writing.

Data processing

Exploratory factor analysis

Exploratory factor analysis was conducted to better analyse the data derived from the computer-based cognitive writing process questionnaire and to determine its validity and reliability. Among the 34 observable variables, five common factors were extracted to confirm the dimensions as well as make an initial test for the construct of the questionnaire. After principal component analysis and varimax orthogonal rotation, five items were deleted, either because their factor loadings

were less than 0.4 or high in two dimensions, or their meanings were hard to interpret or repetitive, so finally 29 items were left. Again, exploratory factor analysis was conducted for the remaining 29 items. The KMO and Bartlett's test indicate that the value of KMO is 0.840, suggesting that these data are fit for factor analysis. The Approx. Chi-Square of Bartlett's Test is 5439.450 and the significance of Bartlett's statistic is 0.000, less than 0.05, proving common factors exist among these items.

According to the information of the rotated component matrix, five factors are identified. Based on the division of Flower and Hayes (1981) for the cognitive writing process, the five factors are respectively named as translating, goal-setting, generating ideas, reviewing and organising ideas (Table 1). The eigenvalue of each factor is above 1 and the cumulative variance of the five factors reaches 58.113%, which explains the variance of the whole questionnaire well. The factor loading of each item can be seen in the range of 0.505 to 0.892, much greater than the acceptable value (0.3). Therefore, it can be concluded that this questionnaire with 29 items has good construct validity. The inner reliability of each factor has been examined and the Cronbach α for translating, goal-setting, generating ideas, reviewing and organising ideas is 0.897, 0.880, 0.805, 0.772 and 0.759 respectively, reaching the measurement standard of 0.70, which shows that each factor has good inner reliability. In addition, the Cronbach α of the total questionnaire is 0.877, showing that this questionnaire has relatively high reliability.

Scoring procedure

The experiments taken by Russell and Tao (2004a, b) confirmed that essays presented in handwritten form received significantly higher scores than the same composition presented in computer-printed form, because errors in computer text were more visible to the raters. In order to avoid such biases in scoring the two tests, after the participants had finished both tests, the authors, with the help of five graduates, typed word for word the 104 paper-written texts titled 'Reading Selectively or

Table 1. Dimensions, factor loadings, eigenvalue, variance and reliability statistics.

Dimension 1 translating		Dimension 2 goal-setting		Dimension 3 generating ideas		Dimension 4 reviewing		Dimension 5 organising ideas	
Item	Loading	Item	Loading	Item	Loading	Item	Loading	Item	Loading
24	.831	1	.892	12	.836	30	.855	18	.791
25	.829	5	.819	10	.824	31	.725	19	.757
28	.806	3	.805	11	.742	33	.705	20	.646
26	.770	7	.776	13	.622	32	.621	15	.597
27	.710	8	.751	14	.535	34	.530	23	.588
29	.677	2	.598	9	.505				
22	.674								
					Eigenvalue				
4.498		3.864		3.155		2.708		2.627	
					Variance				
15.511		13.325		10.880		9.337		9.059	
					Cronbach's alpha				
.897		.880		.805		.772		.759	

Extensively?' and another 104 titled 'Should the University Campus Be Open to Tourists?' onto the computer through the word processor and mixed them respectively with the 104 computer-written texts of the same topic for scoring. All the texts were coded with a series of numbers before removal of the students' names.

Based on the CET-4 holistic grading benchmarks, scoring was conducted by two EFL teachers with online marking experience of CET-4 writing and familiar with the scoring principles of CET-4 writing. The total score of each writing test was 15. The final score of each essay received the average of the two raters. If there was a discrepancy of more than three points in an essay, the essay was graded by a third rater (the expert of the project), whose assessment was accepted as the final score.

Data analysis

The participants' writing scores in different test media and their computer-based cognitive writing process were analysed by using SPSS 17.0.

Paired-sample *t*-tests were employed to analyse the test-takers' overall scores of paper-written compositions and computer-written ones and to check if there were any significant differences in scores between the two test media for students with high, moderate and low levels of computer familiarity respectively. One-way ANOVA was used to analyse if there were any significant differences in the test-takers' cognitive writing process with different levels of computer familiarity.

Results and discussion

Comparison of writing scores between different test media

It can be seen from Table 2 that the test-takers' mean score on the paper-based writing test is 10.01, slightly higher than that on the computer-based writing test (9.52). The result of the paired sample *t*-test indicates the significance value as 0.000, smaller than 0.05, suggesting that there is a significant difference in test-takers' overall scores between computer-based and paper-based writing tests. Therefore, different test media can influence students' performance in writing tests.

To better understand which groups of the test-takers were mostly influenced by the test media or whether the differences in the test-takers' levels in computer familiarity were reflected in their writing test scores, the authors analysed the writing scores of the paper-based and computer-based writing tests of the three groups. As shown in Table 3, for Group 1, the mean score (10.23) of the paper-based writing test is higher than that (9.13) of the computer-based writing test. The significance value of the two-tailed test is 0.000, lower than 0.05, which indicates that there is a significant difference between the paper-based and the computer-based writing scores for test-takers with a low level of computer familiarity.

Table 2. Comparison of the mean scores on paper/computer-based writing tests.

Medium	*N*	Mean	*SD*	*T*	Sig. (2-tailed)
Computer	208	9.52	1.733	-3.698	.000
Paper	208	10.01	1.410		

Table 3. Comparison of the paper/computer-based writing scores for test-takers with different levels of computer familiarity.

Group	Medium	*N*	Mean	*SD*	*T*	Sig. (2-tailed)
1	Paper	58	10.23	1.417	5.662	.000
	Computer	58	9.13	1.536		
2	Paper	87	9.96	1.327	3.502	.175
	Computer	87	9.51	1.786		
3	Paper	63	9.92	1.499	0.854	.395
	Computer	63	9.80	1.632		

As for Group 2, the mean score of their paper-based writing test is 9.96, also higher than 9.51 of their computer-based one. The significance value for this group is 0.175, higher than 0.05, so no significant difference exists in their writing scores in the two test modes.

In regard to Group 3, the mean scores of their essays in the two writing tests are 9.92 and 9.80 respectively. The significance value (0.395) far exceeds 0.05, meaning there is no significant difference.

To sum up, the computer test medium has a strong effect on the test-takers with a low level of computer familiarity. However, for the test-takers with medium and high levels of computer familiarity, the test medium does not have much influence. Therefore, the differences in computer familiarity can be reflected in the test-takers' writing scores.

Comparison of the computer-based cognitive writing process between test-takers with different levels of computer familiarity

From the above factor analysis of the cognitive writing process, five stages were observed: translating, goal-setting, generating ideas, reviewing and organising ideas. Based on the five stages, comparisons were made to address the issue as to whether the test-takers with different levels of computer familiarity perform differently in the cognitive writing process when taking a computer-based writing test.

In this analysis, computer familiarity serves as the grouping variables and the factor score saved in factor analysis as the dependent variables. The results of the test of homogeneity of variances for this stage are presented in Table 4, which includes the results of all five stages. The significance values for goal-setting, generating ideas, organising ideas and translating stages are more than 0.05, indicating that ANOVA can be conducted in the four stages among the groups. As for the

Table 4. The results of the test of homogeneity of variances for all five stages.

	Levene statistic	df1	df2	Sig.
Goal-setting	.873	2	206	.419
Generating ideas	.891	2	206	.411
Organising ideas	1.285	2	206	.278
Translating	.250	2	206	.779
Reviewing	6.607	2	206	.002

Table 5. The results of ANOVA for the former four stages.

		Sum of squares	df	Mean square	*F*	Sig.
Goal-setting	Between groups	.958	2	.479	.478	.621
	Within groups	207.042	206	1.003		
Generating ideas	Between groups	.975	2	.488	.486	.615
	Within groups	207.025	206	1.003		
Organising ideas	Between groups	9.169	2	4.585	4.710	.010
	Within groups	198.878	206	.973		
Translating	Between groups	12.625	2	6.312	6.506	.002
	Within groups	195.375	206	.970		

reviewing stage, the significance value is 0.002, below 0.05, suggesting that non-parametric analysis should be used.

As shown in Table 5, the significance value of the goal-setting stage is 0.621, largely more than 0.05, indicating that the three groups divided by levels of computer familiarity do not have significant difference in this stage. The mean scores for Groups 1, 2 and 3 in this stage are 3.82, 3.97 and 3.99 respectively (see Appendix 3), implying that they carefully read the title and instructions and then thought of the contexts to be written in the writing test. They did not have problems in understanding the title and the instructions. Computer familiarity, therefore, does not exert influence on their goal-setting.

As for the generating idea stage, the significance value is 0.615, largely higher than 0.05, suggesting there is a very high degree of similarity among the test-takers in this stage. The mean scores for Groups 1, 2 and 3 in this stage are 3.03, 3.22 and 3.30 respectively, which proves that students with different levels of computer familiarity admitted that their ideas generated at the beginning of the writing were not perfect, nor necessarily logical or reasonable. However, they believed that they could think out most of the ideas and contexts related to the topic. Therefore, the computer familiarity has no effect on the test-takers' generating ideas. A possible explanation for this is that the test-takers are familiar with the topic and there are other important factors accounting for the generation of ideas, such as background knowledge, writing proficiency etc.

However, the significance values for the organising ideas and translating stages are 0.010 and 0.002 respectively, less than 0.05, suggesting that there is a significant difference among the three groups in the two stages. In order to find out the differences between groups, multiple comparisons were conducted and the results are presented in Table 6.

Table 6. Multiple comparisons for the three groups in organising ideas and translating.

	Organising ideas		Translating	
Groups	Std error	Sig.	Std error	Sig.
1 & 2	.13332425	.033*	.13310131	.022*
1 & 3	.14184849	.029*	.14161129	.002*
2 & 3	.12016117	.892	.11996025	.560

*The mean difference is significant at the 0.05 level.

The significance values between Groups 1 and 2, Groups 1 and 3, and Groups 2 and 3 in organising ideas are observed to be 0.033, 0.029 and 0.892 respectively, representing a significant difference between the test-takers with low computer familiarity and those with medium and high computer familiarity. The comparison of the mean scores of 2.78, 3.71 and 3.86 for Groups 1, 2 and 3 in this stage indicates that the test-takers with low computer familiarity have difficulties in arranging ideas and contexts on the computer quickly and easily. One possible reason is that they are not accustomed to writing on screen, which affects concentration and thinking. However, most of the test-takers with moderate and high computer familiarity reported that they were able to distinguish the importance of ideas and contexts and to remove some ideas when putting them in good order. Therefore, computer familiarity may possibly discourage the test-takers with a low level of computer familiarity but favour those with moderate and high levels of computer familiarity in this stage.

From the multiple comparisons presented in Table 6 and the mean scores of 2.72, 3.65 and 3.77 for Groups 1, 2 and 3 in the translating stage, conclusions similar to those in the organising idea stage can be drawn, namely that computer familiarity may possibly influence the test-takers with a low level of computer familiarity in the translating stage, as most test-takers with medium and high levels of computer familiarity agreed that they could express ideas using appropriate words and correct sentences, but this was not an easy job for the test-takers with a low level of computer familiarity. In addition, most of the test-takers in Groups 2 and 3 reported that they were able to develop paragraphs in a logical order and to connect their ideas smoothly in the whole essay. However, the test-takers of Group 1 generally found it difficult to use appropriate words and sentence structures to express their ideas, as they were not familiar with keyboarding characters on the computer and spent a long time keyboarding. As a result, their concentration was disturbed.

With regard to the last stage, the result of nonparametric analysis is presented in Table 7. The significance value for the three groups in reviewing is 0.126, higher than 0.05, showing that there is no significant difference in reviewing for the three groups. The mean scores for Groups 1, 2 and 3 in this stage are 3.83, 3.97 and 4.03 respectively. The high agreement of the test-takers in reviewing indicates that on the computer, the test-takers can easily conduct all kinds of operations such as correcting mistakes and deleting unnecessary words, sentences and even paragraphs without influencing the cleanness of the writing. As an important stage, revision on the computer favours the test-takers in such functions as copying, pasting, moving and inserting etc. Therefore, computer familiarity has not influenced their reviewing process.

Table 7. Kruskal-Wallis test result in the reviewing stage.

	Computer familiarity	*N*	Mean rank	Chi-square	df	Asymp. Sig.
Reviewing	1	58	166.25	4.140	2	.126
	2	87	177.05			
	3	63	195.49			
	Total	208				

Table 8. The ranking results of the 35th question for the three groups.

	Group 1		Group 2		Group 3	
Item	Weight %	Rank	Weight %	Rank	Weight %	Rank
35–1	15.80	3	18.74	1	10.63	5
35–2	17.50	2	17.37	2	12.72	4
35–3	18.27	1	9.28	6	9.36	6
35–4	8.94	6	15.51	3	16.13	3
35–5	11.61	5	6.64	8	7.55	8
35–6	12.38	4	8.20	7	8.36	7
35–7	7.36	8	13.61	4	18.94	1
35–8	8.13	7	10.65	5	16.31	2

Factors affecting test-takers' performance in the computer-based writing test

The last question in the questionnaire investigated the main factors affecting the test-takers' performance in the computer-based writing test. The test-takers were required to rank the factors based on their own experience, which included both positive factors such as 'easy correction without impacting the cleanness of paper' (35–4), 'easy arrangement of ideas or contents' (35–7) and 'easily discovering grammar, spelling and punctuation errors' (35–8), as well as negative factors such as 'flashing screen impacting thinking' (35–1), 'sound of other test-takers' keyboarding affecting concentration' (35–2), 'unfamiliarity with keyboarding' (35–3), 'being accustomed to outlining with pen and pencil' (35–5) and 'being unaccustomed to keyboarding while organising ideas' (35–6). The ranking results of the eight factors for each group are sorted on the basis of effect degree (Table 8).

The ranking results of Group 1 are 35–3, 35–2, 35–1, 35–6, 35–5, 35–4, 35–8, 35–7, in which all of the negative factors (35–3, 35–2, 35–1, 35–6, 35–5) seem to weigh heavily against the effectiveness of the computer-based writing test, while the positive factors (35–4, 35–8, 35–7) contribute little to it. For Group 3, the negative factors exert a small effect on them; by contrast, the biggest effects are 35–7, 35–8 and 35–4, all being positive factors. And for Group 2, the order of the effect factors from 35–3 to 35–6 is the same as that of Group 3, but the two groups differ widely in that the first two effect factors of one group are just the last two factors of the other group. From the above-ranking results, the conclusion can be drawn that, when writing on the computer, the respondents in Group 1 are more negatively than positively affected, while those in Group 3 get more positive influences than negative ones. As for the respondents in Group 2, both negative and positive factors seem to have impacted them from different aspects.

Discussion

In order to minimise the subjective factors in scoring, the authors adopted the relatively impartial scoring method to draw the scorers' attention more to the quality than the mode of the test. Results reveal that differences indeed exist in test-takers' overall scores between computer-based and paper-based writing tests. But such differences are mainly brought about by their unfamiliarity with computer-based writing, as only the test-takers with a low level of computer familiarity are affected by the computer test medium. In other words, the test-takers with less computer experience are more

disadvantaged by the computer mode. Thus, the differences in computer familiarity may possibly be reflected in their writing test scores.

To answer the second research question, the authors analysed the five cognitive writing stages. The results indicate that there are similarities and differences in the cognitive writing process among the test-takers with different levels of computer familiarity. Their similarities are in the stages of goal-setting, generating ideas and reviewing, while their differences are in organising ideas and translating stages. In other words, the computer familiarity may influence the test-takers with a low level of computer familiarity in terms of organising ideas and translating. However, the ranking results show that different aspects exert different degrees of effects on different test-takers when writing on the computer. For the test-takers with less computer familiarity, the negative factors are mainly 'unfamiliarity with keyboarding', 'flashing screen impacting thinking' and 'sound of other test-takers' keyboarding affecting concentration'.

Conclusion

It has been established in the present study that computer familiarity is one of the main factors associated with the test-takers' writing scores and the test-takers with different levels of computer familiarity perform differently in the cognitive writing process, owing to the nature of the computer test medium itself. Considering the facts that computer-based language tests have become more prevalent and important in language instruction and many international testing services have begun to use computerised tests in their programmes, in order to minimise the adverse effects, language teachers are advised to offer related training before implementing the new test model. In addition, teachers are urged to integrate the use of computers into their teaching of writing, to make use of the advantages of computers to develop a composition e-portfolio to improve students' writing ability, meanwhile, to make them increasingly familiar with computer-based testing. To let test-takers bring their ability into full play in computer-based writing tests, test services are also recommended to provide different test media for test-takers to choose, especially in large-scale assessment.

The findings of this study do bring some useful information. However, there still are some limitations embedded in it. Firstly, owing to the limitation of time and space, the research participants were from only one university. The conclusion of this study can only be generalised in this university or for other universities of a similar nature. Therefore, the application of the results into other situations should be considered carefully. Secondly, the data of this study were mainly collected with questionnaire surveys, which could not investigate the in-depth and detailed information of the factors affecting the performance of different test-takers between computer-based and paper-based writing tests. More full and accurate evidence is needed to prove the effects of test media on test-takers' performance. Therefore, future research should consider the sampling from different universities or colleges. Besides grouping the sampling by levels of computer familiarity, one can also group them by gender, geographic location, computer anxiety and other specific grouping variables. In addition to argumentation, other writing styles such as narration, description and exposition can be used in writing tests to see whether the test-takers perform differently in different test media. Thirdly, more detailed analysis that presents the test-takers' differences in grammar, sentence structure and content etc. can

be taken into account. Fourthly, qualitative research, especially micro-research aimed at individual cases, should never be neglected in investigating test-takers' detailed cognitive writing process. Lastly, as this research only compared the cognitive writing process of the test-takers with different levels of computer familiarity on a computer-based writing test, the comparison of the cognitive writing process between computer-based and paper-based tests could be another research topic in future.

All in all, although computers have been widely used in language tests abroad, they are still not very extensively applied in China. Therefore, more empirical and theoretical studies are needed to make the application of computers more efficient in the reform of language testing in China.

Funding

This work was supported by the Foundation of Education Committee of Chongqing, China [Grant No. 101303].

Notes on contributors

Xiao-Ling Zou is a Professor in the School of Foreign Languages and Cultures, and a researcher in Research Academia for Linguistics, Cognition & Application, Chongqing University, China. She is the author of over 30 published papers, director of over 10 research projects, chief or vice chief editor of more than 40 College English textbooks, one of the main translators of the Oxford Advanced Learner's English–Chinese Dictionary (the 6th, 7th and 8th editions) and one of the main lexicographers of the Oxford Chinese Dictionary. Her research interests are computer-assisted language learning, English instruction and bilingual dictionary research.

Yan-Min Chen is a graduate in the School of Foreign Languages and Cultures, Chongqing University, China. Her main research interest is computer-assisted language learning.

References

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Lawrence Erlbaum.

Breland, H., Lee, Y.-W., & Muraki, E. (2005). Comparability of TOEFL CBT essay prompts: Response-mode analyses. *Educational and Psychological Measurement, 65*, 577–595.

Bridgeman, B., & Cooper, P. (1998, April). *Comparability of scores on word-processed and handwritten essays on the graduate management admissions test*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Bridwell, L. S., Sirc, G., & Brooke, R. (1985). Revising and computing: Case studies of student writers. In S. W. Freedman (Ed.), *The acquisition of written language: Response and revision* (pp. 172–194). Norwood, NJ: Ablex.

Celce-Murcia, M., & Olshtain, E. (2000). *Discourse and context in language teaching: A guide for language teachers*. Cambridge: Cambridge University Press.

Chen, J., White, S., McCloskey, M., Soroui, J., & Chun, Y. (2011). Effects of computer versus paper administration of an adult functional writing assessment. *Assessing Writing, 16*, 49–71.

Chen, K. T. (2012). Elementary EFL teachers' computer phobia and computer self-efficacy in Taiwan. *TOJET: The Turkish Online Journal of Educational Technology*, *11*(2). Retrieved from http://www.tojot.net/articles/v11i2/11210.pdf

Chua, Y. P., & Don, Z. M. (2013). Effects of computer-based educational achievement test on test performance and test takers' motivation. *Computers in Human Behavior, 29*, 1889–1895.

Flower, L., & Hayes, R. J. (1981). A cognitive process theory of writing. *College Composition and Communication, 32*, 365–387.

Fulcher, G. (1999). Computerizing an English language placement test. *ELT Journal, 53*, 289–299.
Gao, B. L. (2007). A comparative study of COPT and DOPT. *Computer-Assisted Foreign Language Education, 2*, 73–76.
Genc, H. (2012). An evaluation study of a CALL application: With BELT or without BELT. *TOJET: The Turkish Online Journal of Educational Technology, 11*, 44–54.
Goldberg, A., Russell, M., & Cook, A. (2003). The effect of computers on student writing: A meta-analysis of studies from 1992 to 2002. *Journal of Technology, Learning, and Assessment, 2*, 2–51.
Grabe, W. (2001). Notes toward a theory of second language writing. In T. Silva & P. K. Matsuda (Eds.), *On second language writing* (pp. 39–57). Mahwah, NJ: Lawrence Erlbaum.
Grabe, W., & Kaplan, R. B. (1996). *Theory and practice of writing: An applied linguistic perspective*. London: Longman.
Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy & S. E. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 1–27). Mahwah, NJ: Lawrence Erlbaum.
He, Q., & Tymms, P. (2005). A computer-assisted test design and diagnosis system for use by classroom teachers. *Journal of Computer Assisted Learning, 21*, 419–429.
Horkay, N., Bennett, R. E., Allen, N., Kaplan, B., & Yan, F. (2006). Does it matter if I take my writing test on computer? An empirical study of mode effects in NAEP. *Journal of Technology, Learning, and Assessment*, *5*(2). Retrieved from http://www.jtla.org
Hsiao, H.-C., Tu, Y.-L., & Chung, H.-N. (2012). Perceived social supports, computer self-efficacy, and computer use among high school students. *TOJET: The Turkish Online Journal of Educational Technology, 11*, 167–177.
Imhof, M., Vollmeyer, R., & Beierlein, C. (2007). Computer use and the gender gap: The issue of access, use, motivation, and performance. *Computers in Human Behavior, 23*, 2823–2837.
Jin, Y., & Wu, J. (2010). A preliminary study of the validity of the Internet-based CET: Factors affecting test-takers' perception of and performance on the test. *Computer-Assisted Foreign Language Education, 2*, 3–10.
Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. E. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 57–71). Mahwah, NJ: Lawrence Erlbaum.
Lee, H. K. (2004). A comparative study of ESL writers' performance in a paper-based and a computer-delivered writing test. *Assessing Writing, 9*, 4–26.
Lee, Y.-J. (2002). A comparison of composing processes and written products in timed-essay tests across paper-and-pencil and computer modes. *Assessing Writing, 8*, 135–157.
Leeson, H. V. (2006). The mode effect: A literature review of human and technological issues in computerized testing. *International Journal of Testing, 6*, 1–24.
Li, Q. H. (2006). Equivalence studies of paper-and-pencil based language testing and computer based language testing: A survey. *Foreign Language World, 4*, 73–78.
Li, Q. H., & Kong, W. (2009). Computer-based language testing and its validation. *Foreign Language World, 3*, 66–72.
MacArthur, C. A. (1988). The impact of computers on the writing process. *Exceptional Children, 54*, 536–542.
MacCann, R., Eastment, B., & Pickering, S. (2002). Responding to free response examination questions: Computer versus pen and paper. *British Journal of Educational Technology, 33*, 173–188.
McCutchen, D. (1986). Doman knowledge and linguistic knowledge in the development of writing ability. *Journal of Memory and Language, 25*, 431–444.
McDonald, A. S. (2002). The impact of individual differences on the equivalence of computer-based and paper-and-pencil educational assessments. *Computers & Education, 39*, 299–312.
Neuwirth, C. M. (1990). *Does word processing improve students' writing? A critical appraisal and assessment*. (Final Report to FIPSE). ERIC Document Reproduction Service No. ED329994.
Norris, J. T., Pauli, R., & Bray, D. E. (2007). Mood change and computer anxiety: A comparison between computerised and paper measures of negative affect. *Computers in Human Behavior, 23*, 2875–2887.

Organisation for Economic Co-operation and Development. (2010). *PISA computer-based assessment of student skills in science*. Retrieved from http://www.oecd.org/publishing/corrigenda

Owston, R. D., Murphy, S., & Wideman, H. H. (1992). The effects of word processing on students' writing quality and revision strategies. *Research in the Teaching of English, 26*, 249–276.

Piaw, C. Y. (2012). Replacing paper-based testing with computer-based testing in assessment: Are we doing wrong? *Procedia – Social and Behavioral Sciences*, *64*, 655–664.

Powers, D. E., Fowles, M. E., Farnum, M., & Ramsey, P. (1994). Will they think less of my handwritten essay if others word process theirs? Effects on essay scores of intermingling handwritten and word-processed essays. *Journal of Education Measurement, 31*, 220–233.

Rohman, D. G. (1965). Pre-writing: The stage of discovery in the writing process. *College Composition and Communication, 16*, 106–112.

Russell, M. (1999). Testing on computers: A follow-up study comparing performance on computer and on paper. *Education Policy Analysis Archives*, *7*(20), 1–47. Retrieved from http://epaa.asu.edu/epaa/v7n20/

Russell, M., & Haney, W. (1997). Testing writing on computers: An experiment comparing student performance on tests conducted via computer and via paper-and-pencil. *Educational Policy Analysis Archives, 5*(3), 1–20.

Russell, M., & Plati, T. (2000). Effects of computer versus paper administration of a state-mandated writing assessment. *Teachers College Record*. Retrieved from http://www.bc.edu/research/intasc/PDF/ComputerVsPaperStateWriting.pdf

Russell, M., & Tao, W. (2004a). Effects of handwriting and computer-print on composition scores: A follow-up to Powers, Fowles, Farnum, & Ramsey. *Practical Assessment, Research & Evaluation*, *9*(1). Retrieved from http://PAREonline.net/getvn.asp?v=9&n=1

Russell, M., & Tao, W. (2004b). The influence of computer-print on rater scores. *Practical Assessment, Research & Evaluation*, *9*(10). Retrieved from http://PAREonline.net/getvn.asp?v=9&n=10

Scheuermann, F., & Pereira, A. G. (2008). *Towards a research agenda on computer-based assessment: Challenges and needs for European Educational Measurement* (JRC Scientific and Technical Reports). Luxembourg: Office for Official Publications of the European Communities.

Sommers, N. (1980). Revision strategies of student writers and experienced adult writers. *College Composition and Communication, 31*, 378–388.

Sun, Y. C. (2007). A contrast analysis and study of anxiety in test of spoken English between computer-aided and traditional form. *Journal of Changchun Normal University* (Humanities and Social Sciences)*, 4*, 23–26.

Taylor, C., Jamieson, J., Eignor, D., & Kirsch, I. (1998). *The relationship between computer familiarity and performance on computer-based TOEFL test tasks* (TOEFL Research Reports 61: ETS Research Report 98–8). Princeton, NJ: Educational Testing Service.

Taylor, C., Kirsch, I., Eignor, D., & Jamieson, J. (1999). Examining the relationship between computer familiarity and performance on computer-based language tasks. *Language Learning, 49*, 219–274.

Vispoel, W. P., Rocklin, T. R., & Wang, T. (1994). Individual differences and test administration procedures: A comparison of fixed-item, computerized-adaptive, and self-adapted testing. *Applied Measurement in Education, 7*, 53–79.

Weir, C., O'Sullivan, B., Jin, Y., & Bax, S. (2007). Does the computer make a difference? The reaction of candidates to a computer-based versus a traditional hand-written form of the IELTS writing component: Effects and impact. *IELTS Research Reports, 7*, 1–37.

Whithaus, C., Harrison, S. B., & Midyette, J. (2008). Keyboarding compared with handwriting on a high-stakes writing assessment: Student choice of composing medium, raters' perceptions, and text quality. *Assessing Writing, 13*, 4–25.

Wolfe, E. W., & Manalo, J. R. (2004). Composition medium comparability in a direct writing assessment of non-native English speakers. *Language Learning & Technology, 8*, 53–65.

Zou, X. L., & Zhang, X. N. (2013). Effect of different score reports of Web-based formative test on students' self-regulated learning. *Computers & Education*, *66*, 54–63.

Appendix 1. Computer familiarity questionnaire

The questionnaire is designed for research on your computer familiarity. Please answer each of the questions honestly and frankly according to your own experience or opinions. There are no 'correct' or 'incorrect' answers. All the data collected are anonymous and confidential and will be seriously considered and used for research only. Thank you for your cooperation and patience.

Part I Personal Information

Student No.:_____ sex: _____ grade:_____ major: _____

Part II Computer familiarity

1	How comfortable are you with using the computer? A. Not at all comfortable B. Somewhat comfortable C. Comfortable D. Very comfortable
2	How comfortable are you with using the mouse? A. Not at all comfortable B. Somewhat comfortable C. Comfortable D. Very comfortable
3	How comfortable are you with using a computer to write a paper? A. Not at all comfortable B. Somewhat comfortable C. Comfortable D. Very comfortable
4	How comfortable would you be taking a writing test on a computer? A. Not at all comfortable B. Somewhat comfortable C. Comfortable D. Very comfortable
5	How would you rate your ability to use a computer? A. Poor B. Fair C. Good D. Excellent
6	How often do you use a computer? A. Less than once a week B. 1–2 times a week C. 3–5 times a week D. Once or more a day
7	How often do you use the internet? A. Less than once a week B. 1–2 times a week C. 3–5 times a week D. Once or more a day
8	How often do you use a computer to send or receive an email? A. Less than once a month B. Less than once a week C. Less than once a day D. Once or more a day
9	How often do you use word-processing in English? A. Less than once a month B. Less than once a week C. Less than once a day D. Once or more a day
10	How often do you use word-processing in your native language? A. Less than once a month B. Less than once a week C. Less than once a day D. Once or more a day
11	How often do you use Microsoft Office Excel, PowerPoint or graphics? A. Less than once a month B. Less than once a week C. Less than once a day D. Once or more a day

Appendix 2. Computer-based cognitive writing process questionnaire

The questionnaire is designed for research on your computer-based writing test processes. Please answer each of the questions honestly and frankly according to your own experience or opinions. There are no 'correct' or 'incorrect' answers. All the data collected are anonymous and confidential and will be seriously considered and used for research only. Thank you for your cooperation and patience.

Part I Personal Information

Student No.:_____ sex:_____ grade:_____ major:_____

Part II Computer-based Cognitive writing processes

According to your experience and understanding of your writing on the computer, please choose just one closest answer and put a circle (○, for example ①) in the column that corresponds to your opinion.

1 = strongly disagree; 2 = disagree; 3 = neutral; 4 = agree; 5 = strongly agree

01	I considered seriously the meaning of each word in the title for accurate understanding.	1	2	3	4	5
02	I thought of what I was required to write after reading the title and instructions.	1	2	3	4	5
03	I thought of how to write my answer so that it would respond well to the title.	1	2	3	4	5
04	I thought of how to satisfy readers or examiners.	1	2	3	4	5
05	I was able to understand the instructions for this writing test completely.	1	2	3	4	5
06	I know a lot about this topic and have enough ideas to write about this topic.	1	2	3	4	5
07	I was able to understand the title quickly and grasp the important information in it.	1	2	3	4	5
08	I felt it was easy to produce enough information related to the essay from memory.	1	2	3	4	5
09	I planned an outline on paper before starting to write on computer.	1	2	3	4	5
10	Ideas occurring to me at the beginning tended to be complete.	1	2	3	4	5
11	Ideas occurring to me at the beginning were logically organised.	1	2	3	4	5
12	I spent less time organising ideas on computer-based than paper-based writing.	1	2	3	4	5
13	I thought of most of the ideas related to the essay before planning an outline.	1	2	3	4	5
14	I thought of most of my ideas related to the essay while planning an outline.	1	2	3	4	5
15	I was able to prioritise the ideas quickly.	1	2	3	4	5
16	I thought about and adjusted the content of the essay while I was writing.	1	2	3	4	5
17	Some ideas had to be removed while I was putting them in good order.	1	2	3	4	5
18	I wrote down the whole sentence after I constructed it in my mind.	1	2	3	4	5
19	I was able to distinguish between primary and secondary ideas or contents.	1	2	3	4	5
20	I was able to put my ideas or content in good order.	1	2	3	4	5
21	I felt it was easy to put my ideas or content in good order.	1	2	3	4	5
22	I thought of most of my ideas for the essay while I was actually writing it.	1	2	3	4	5
23	I organised the structure of the whole essay in my mind before actually writing it.	1	2	3	4	5

(*Continued*)

Appendix 2. (*Continued*)

24	I was able to express my ideas by using appropriate words.	1	2	3	4	5
25	I was able to express my ideas using correct sentence structure.	1	2	3	4	5
26	I was able to develop any paragraph by putting sentences in logical order.	1	2	3	4	5
27	I felt it was easy to express ideas using the appropriate words.	1	2	3	4	5
28	I felt it was easy to express ideas using the correct sentences.	1	2	3	4	5
29	I was able to connect my ideas smoothly in the whole essay.	1	2	3	4	5
30	I reviewed the correctness of the contents and their order while writing this essay.	1	2	3	4	5
31	I reviewed the correctness of the contents and their order after finishing this essay.	1	2	3	4	5
32	I reviewed the appropriateness of the words while writing and after finishing this essay.	1	2	3	4	5
33	I felt it was easy to review or revise the whole essay.	1	2	3	4	5
34	I felt it was easy to review or revise the paragraphs and sentences.	1	2	3	4	5
35	Sort the factors affecting your performance in computer-based writing test in order of effect degree from high to low based on your own experiences. (1) I feel uncomfortable because the flashing screen impacts my thinking. (2) Other test-takers' keyboard tapping affects my concentration. (3) My expression ability has been limited for lack of the skill in typing. (4) Revision on the computer is easier and can keep the page neat and clean. (5) My writing time lengthens because I am accustomed to outlining with pen-and-pencil first. (6) I'm not accustomed to keyboarding while organising ideas. (7) It's easier to arrange ideas or contents of the essay on computer. (8) It's easier to discover grammar, spelling and punctuation errors on the computer.					

Appendix 3. Mean scores of test-takers with different levels of computer familiarity in five writing sub-processes

Mean scores for test-takers with low computer familiarity

Item	Translating	Item	Goal-setting	Item	Generating ideas	Item	Reviewing	Item	Organising ideas
24	2.72	1	3.99	12	2.87	30	4.08	18	3.54
25	2.61	5	3.73	10	2.13	31	4.06	19	2.18
28	2.33	3	4.01	11	2.27	33	3.52	20	2.55
26	2.31	7	3.51	13	3.61	32	3.94	15	2.11
27	2.78	8	3.58	14	3.58	34	3.56	23	3.52
29	2.71	2	4.1	9	3.74				
22	3.58								
Mean	2.72		3.82		3.03		3.832		2.78

Mean scores for test-takers with medium computer familiarity

Item	Translating	Item	Goal-setting	Item	Generating ideas	Item	Reviewing	Item	Organising ideas
24	3.58	1	4.02	12	3.16	30	4.07	18	3.77
25	3.57	5	3.97	10	2.28	31	4.14	19	3.63
28	3.84	3	4.08	11	2.7	33	3.77	20	3.76
26	3.55	7	3.73	13	3.88	32	3.95	15	3.68
27	3.76	8	3.87	14	3.72	34	3.91	23	3.7
29	3.67	2	4.12	9	3.57				
22	3.61								
Mean	3.65		3.965		3.22		3.968		3.708

Mean scores for test-takers with high computer familiarity

Item	Translating	Item	Goal-setting	Item	Generating ideas	Item	Reviewing	Item	Organising ideas
24	3.72	1	4.07	12	3.2	30	4.16	18	3.8
25	3.59	5	4.05	10	2.57	31	4.17	19	3.83
28	3.85	3	4.1	11	2.81	33	3.83	20	3.94
26	3.61	7	3.81	13	3.94	32	4.08	15	3.95
27	3.88	8	3.73	14	3.88	34	3.93	23	3.76
29	3.8	2	4.16	9	3.41				
22	3.96								
Mean	3.77		3.987		3.302		4.034		3.856