

Using Information and Communication Technology for the
Assessment of Writing in Chinese Language

Esther Yee

Louis Yim

Chua Lay Keng

Lyndon Lim

Singapore Examinations and Assessment Board, Singapore

Abstract

Singapore schools have leveraged on the affordances of Information and Communication Technology (ICT) to develop students into proficient users of their Mother Tongue Languages (MTL) who can communicate confidently and effectively in real-life situations. In particular, computer-based writing has been introduced in schools as part of the MTL curriculum to prepare students for the future where computer-based writing is fast becoming a norm in the workplace and in social communications. Where appropriate and for better alignment with the teaching and learning of the curriculum, computer-based writing would be planned and introduced in the assessment of written interaction skills.

This paper shares some key findings on the use of a computer-based writing platform for high-school students taking Chinese Language B Syllabus to compose email responses and blog entries. The quality of students' written responses from two separate studies: a) between computer- and paper-based writing; and b) between progressive stages of the computer-based writing, were analysed. Students' and teachers' feedback about computer-based versus paper-based writing tasks are also discussed in this paper.

Keywords: language testing, information and communications technology, writing test

Using Information and Communication Technology for the Assessment of Writing in Chinese Language

Since the 1990s, there has been a shift towards using computers for testing and measurement purposes. This shift has been brought about by the advantages of computer-based (CBT) testing over paper-based testing (PBT) for certain modes and purposes of tests. As summarised by Bridgeman (2008), these advantages include:

“paperless test distribution and data collection, greater standardization of test administrations, monitoring of student motivation, obtaining machine-scorable responses for writing and speaking, providing standardized tools for examinees (e.g., calculators and dictionaries), and the opportunity for more interactive question types” (p.39).

Further to these advantages of CBT, new constructs that are important to thrive in the 21st century but have been difficult to assess via PBT can now be realised. In Singapore, schools have leveraged on the affordances of Information and Communication Technology (ICT) to develop students into proficient users of their Mother Tongue Languages¹ (MTL) who can communicate confidently and effectively in real-life situations. In particular, the Ministry of Education, Singapore (MOE, 2012) has recently launched the i-MTL portal in all schools across all levels (Grade 4 to Grade 11) to support the teaching and learning of computer-based writing in MTL. Students could enhance their written communication skills through the interactivity and multi-modal feedback features available in the portal.

Literature Review

While CBT may assess what PBT is unable to and provide other advantages and possibilities, one of the main concerns is the issue of the equivalence of CBT and PBT. According to Ford, Vitelli, and Stuckless (1996), CBT and PBT which are intended to be interchangeable should have their equivalence empirically verified and not simply assumed before validity tests can be performed on the computerised tests. To compare outcomes of the two tests, as cited in Kveřton, Jelı́nek, Voborřil, and Klimusova (2007), APA (1986) recommended the following psychometric properties be studied: “(1) descriptive statistics: means, variances, distributions, and rank orders of scores; (2) construct validity; (3) reliability” (p.33). Xie (2001) detailed the process of students using the keyboard (as opposed to handwriting recognition) for Chinese writing into (1) thinking of the characters and how they sound, (2) spelling them using hanyu pinyin rules, (3) typing the letters using the keyboard, (4) distinguishing among the characters listed by the software and (5) selecting the desired characters from the list. Ardila (2013) argued that handwriting demands significantly different cognitive and motor abilities than keyboarding tasks. Nevertheless, he went on to cite a study (Lin, Xiao, Shen, Zhang & Weng, 2007) which found, through fMRI technology that brain activity of students writing with Chinese characters and *hanyu pinyin* activated essentially the same brain region. It can therefore be surmised that CBT and PBT are not necessarily assessing the same construct but there could be a substantive overlap of basic core writing skills.

In Singapore schools, a change in assessment format, including a change in the mode of assessing writing to one which is computer-based, will always be preceded by a timely change in the teaching and learning syllabus. CBT and PBT are therefore not concurrently offered to students of the same cohort. To transition from PBT to CBT, the premise is that it is already in

¹ These are Chinese Language, Malay Language and Tamil Language

the curriculum and there is a need for alignment with the curriculum, i.e. for students to be assessed the way they have been taught. The question is whether students find the CBT as accessible as any other examination. Indicators of accessibility could include quantity as well as quality content of the composed responses, appropriate language use and sustained writing which would be evidenced in the response length.

In a review of previous research, Bugbee (1996) found that even if identical tests are given, different modes of administration matters. While some studies yield better results in CBT, others yield the contrary. It appeared that content familiarity was the one factor that was most strongly associated with the mode effect. With the advent of computers in this period and the drive by the Singapore Ministry of Education to use more computer related pedagogies in teaching and learning, students in Singapore have been taught basic keyboarding skills in the respective MTL from Grades 3 or 4. Thus, it may be assumed that students have sufficient computer and content familiarity. However, studies must still be conducted to demonstrate that the mode effect does not contribute to any sources of irrelevant variance and consequently undermine the validity of test scores.

Complementary to psychometric factors and the mode effect that define the equivalence of CBT to PBT are factors related to test construction and impact to the student directly. These factors are critical as they define the level of student acceptability in CBT. As Terzis and Economides (2011) posited, “the effective development of a computer based assessment depends on students’ acceptance” (p.1032). Choi, Kim and Boo (2003) highlighted the following factors that will directly affect the student:

- prior exposure to computers (computer familiarity) and computer anxiety factors;
- the task types of multiple choice items;
- the existence of graphic information;
- the inclusion of lengthy reading passages;
- speededness/time constraints; and
- the type of computer interface (p.297).

Results from previous studies to show that students accept and prefer CBT over PBT have been inconclusive. Terzis and Economides (2011) highlighted that in their review of previous studies, students prefer CBT to PBT as CBT is “more promising, credible, objective, fair, interesting, fun, fast and less difficult or stressful” (p.1032). Though there is a perceived ease of use in CBT, some students are also apprehensive over technological issues such as test security and the possibility of cheating in CBT. In their research on implementing computer-based assessment (CBA) in higher education, Deutsch, Herrmann, Frese, and Sandholzer (2012) concluded that while the CBA experience had “a positive influence on attitudes towards CBA, there were strong reservations about technical problems influencing the test performance when used for summative assessment” (p.1068). With these findings, it is pertinent to gather students’ feedback on CBT and also to convey assurances to them.

Research Questions

Two studies were conducted to investigate the performance of two groups of Grade 11 students taking the Chinese Language (CL) B subject using computer text entry for the writing tasks. The following research questions were for investigation:

- (i) *Study I:* For a group of students familiar with pen-based writing and given sufficient exposure to computer-based writing, are the two modes of testing equally accessible judging from the difference, if any, in the quality of writing task responses?

- (ii) *Study II:* For a group of students familiar with computer-based writing, is there any progression in quality when students are tested over a three-month period?

The writing tasks for both groups of students were identical in format though presented in the respective modes of delivery. Students could choose to either write a response to an incoming email or create a blog entry within a duration of 50 minutes.

Students' responses were analysed quantitatively (in terms of raw marks and length of response) and qualitatively (in terms of vocabularies, language use and organisation) based on the professional judgment of experienced markers. Raw marks were awarded based on a rubric where a total of 20 marks were allocated equally between Content and Language Use.

Study I: Comparison between PBT and CBT

A sample of about 100 Grade 11 students participated in this study. The sample was drawn from 12 Junior Colleges (JCs) and is representative in term of the range of ability among students taking the subject. These students had been taught email writing or blog entry. They had also been exposed to computer-based writing through school or take-home assignments. Two writing tasks were developed for the study and these were judged to be of equivalent difficulty by assessment specialists. One was packaged as a computer-based test (CBT), which was administered onscreen on a computer via a proprietary e-platform. The other writing task was administered as a paper-based test (PBT). Students took both PBT and CBT with a short break between tests. The sequence in which they took the test was randomly assigned to them, with half the students taking PBT first and the other half taking CBT first. At the end of the tests, a survey was administered to the students to gather feedback on their test-taking experience.

Results

The *raw mark* and *word count* of the responses were analysed and the descriptive statistics shown in Table 1 below. The correlations of each pair of variables between modes and t-values evaluating the statistical significance of the differences between means are also provided.

Table 1
Comparison of Statistical Results between PBT and CBT Modes for the Variables of Raw Mark and Word Count (n=92)

Variable	Descriptive statistics Mean (s.d.)		n	correlation	t-value (p)
	CBT	PBT			
Raw Mark	9.44 (4.04)	9.42 (3.96)	92	0.49	0.15 (p>0.05)
Word Count	317 (97.6)	246 (74.7)	92	0.52	7.60 (p<0.0001)

Between PBT and CBT, students' performance in terms of *raw mark* is comparable. The t-test results show that the difference in mean *raw mark* is not statistically significant (p>0.05). This suggests the mode of testing did not impinge on students' ability to demonstrate

their writing skills and the CBT is equally accessible as PBT. The moderate correlation between modes also suggests some substantial overlap of basic writing skills.

For *word count*, there is a large difference in the means of about 70 characters between PBT and CBT. The t-test results indicate it is statistically significant ($p < 0.05$). This suggests that students were motivated to write more in CBT mode than in PBT mode as evident from the longer texts they produced under CBT mode.

Feedback from participants

Students' views on computer-based writing gathered through a survey were favourable. The response rates to two of the survey questions are shown below.

<i>Question 1: Compared to writing on paper, was typing on a computer more or less enjoyable?</i>		
Typing was more enjoyable	Typing was about as enjoyable as writing	Writing was more enjoyable
62 (67.4%)	16 (17.4%)	14 (15.2%)

<i>Question 2: Do you look forward to using a computer for CL examination?</i>	
Yes	No
76 (82.6%)	16 (17.4%)

About 85% students responded that they found typing about as enjoyable, if not more so, than writing. About 83% students looked forward to using a computer for examination purposes. Overall, their experience in CBT was positive and anecdotal comments from them included

- “more time to plan”
- “it’s faster than writing”
- “love the easiness of editing on computer”

Comments from teachers included

- “introduction of ICT in examination is timely”
- “it is vital to equip our students with the 21st Century Competencies skills to enable them to thrive in the highly connected world”

Study II: Comparison of writing response over two CBT sessions

A representative sample of 40 Grade 11 students from 7 Junior Colleges participated in this study. Their responses to two computer-based writing tasks, administered as CBTs under school-based examination conditions in July and September, were analysed and compared. These two writing tasks were judged to be of equivalent difficulty by assessment specialists.

Results

The results of two tests are presented in Table 2 below.

Table 2

Comparison of Statistical Results for Students in Two Tests for the Variables of Raw Mark and Word Count (n=40)

Variable	Descriptive statistics Mean (s.d.)		n	correlation	t-value (p)
	Test 1	Test 2			
Raw Mark	11.53 (2.88)	10.98 (2.99)	40	0.632	1.38 (p>0.05)
Word Count	346.4 (92.5)	373.05 (121.6)	40	0.742	-2.068 (p<0.05)

Although there is a slight decrease in the mean mark between Test 1 and Test 2, this difference is not statistically significant ($p>0.05$). This could also suggest that students were equally well-prepared for both tests.

However, there is an increase in the mean of *word count*, and such an increase does not happen by chance as the t-test result shows that the means are significantly different between the two tests. This could be due to candidates making an effort to write more in Test 2. It is interesting to note that the increase in *word count* did not carry with it any significant increase in the mean *raw mark*. This could suggest students are nearing or at their peak performance by the time the tests were conducted. The correlation in terms of *raw mark* and *word count* are in the region of 0.6-0.7, indicating the two variables correlate moderately between the two tests.

Script quality in CBT mode

The answer scripts from 20 students were selected randomly for qualitative analysis in terms of their content, language use and organization. In general, all 20 students were able to provide some personal responses based on the stimulus provided. Majority of them were able to develop some relevant ideas in a fairly clear and coherent manner and were able to meet the requirement of at least 200 characters for the e-mail or blog. Many of them produced simple sentence patterns and used basic, though sometimes inappropriate, vocabularies in their writing.

The students showed slight improvement in their writing as they progressed from Test 1 to Test 2. Generally, students were well prepared and had the essential skills required for the computer-based writing tasks. There were few major errors in students' responses. The use of vocabulary was perhaps limited but the meaning of the message was clear. In terms of language use and organization (such as sentence structure, grammar and word usage), all tests have the following characteristics:

- fairly well done, though with rambling and rather repetitive sentence structures;
- succinct, but with the use of limited vocabularies;
- flat content.

Comparing with Test 1, there were more students using more complex vocabularies (such as “动脑筋”、“辅导课”、“无聊”) in their writing during Test 2. Sentences were also longer and more complex in structure (such as: “我们的心里肯定会感觉甜滋滋的，像吃了

蜜糖似的”，“每一年都能看到不同的学生表演，有的唱歌，有的跳舞，多精彩啊!"). Majority of the students were able to develop ideas relevant to the given topics in a fairly clear and coherent manner.

In computer-based writing in CL, students need to activate both the *pinyin* strategies and ‘character recognition’ strategies which are taught as part of the CL curriculum. Analysis of students’ scripts suggested that a number of them had either failed to key-in the right *pinyin* which led to the wrong range of characters available for selection or they were unable to identify the desired characters from the selection list even when the correct *pinyin* was used. Some examples are illustrated as follows:

Category of error: correct pinyin but wrong selection

Correct <i>Pinyin</i>	Intended (inferred from context)	Wrong Selection
jiangtang	讲堂	将堂
fengherili	风和日丽	风和日力
meinian	每年	没年
xianzai	现在	先在

Category of error: wrong pinyin entered

Intended (inferred from context)	Wrong <i>Pinyin</i>
开心 kaixin	开兴 kaixing
因为 yinwei	应为 yingwei
重要 zhongyao	总要 zongyao
怎么 zenme	这么 zheme

Conclusion

Findings from this study showed that the overall performance of students in computer-based writing and pen-based writing (as measured by *raw mark*) was comparable. The study also showed that students under CBT mode tend to produce longer writing tasks as compared to PBT mode. Using *word count* as a proxy of the willingness to write in CL, this study further suggests that the use of computer has, to some extent, helped motivate students to express themselves in writing.

Findings also revealed that the *mean mark* for computer-based writing tested over a three-month period was comparable ($p>0.05$). This suggests that students were equally well-prepared between the two tests. There is, however, an increase in the mean *word count* which is not due to chance. This could be due to candidates making an effort to write more at the second test.

Qualitative analysis of writing showed that students were well-prepared with the essential skills required for computer-based writing. They were able to write coherently and could develop their ideas according to the topics. Nevertheless there is room for them to strengthen their skills in *pinyin* and character recognition.

Collectively, the two studies suggested empirically that students felt writing using computer text input was as accessible as the pen-and-paper mode. It is nevertheless crucial to familiarise students with the testing conditions and content when transitioning from PBT to CBT. With sufficient practice, the testing experience under CBT would become more positive and even enjoyable. More importantly, students would be future-ready and better equipped for the working world where computer-based writing is gaining importance.

References

- Ardila, A. (2012). Neuropsychology of Writing. In Grigorenko, E.L., Mambrino, E., & Preiss, D. D. (Eds), *Writing: A Mosaic of New Perspectives* (317-318).
- Bridgeman, B. (2008). Experiences from Large-Scale Computer-Based Testing in the USA. In Friedrich, S., & Bjornsson, J. (Eds.), *The transition to computer-based assessment - new approaches to skills assessment and implications for large-scale testing* (39-44). Retrieved from: <http://www.gesci.org/assets/files/reporttransition.pdf>.
- Bugbee Jr., A. C. (1996). The equivalence of paper-and-pencil and computer-based testing. *Journal Of Research On Computing In Education*, 28(3), 282.
- Choi, I. C., Kim, K. S., & Boo, J. (2003). Comparability of a paper-based language test and a computer-based language test. *Language Testing*, 20(3), 295–320.
- Deutsch, T., Herrmann, K., Frese, T., Sandholzer, H. (2012). Implementing computer-based assessment – A web-based mock examination changes attitudes. *Computers & Education*, 58, 1068–1075.
- Ford, B. D., Vitelli, R., & Stuckless, N. (1996). The effects of computer versus paper-and-pencil administration on measures of anger and revenge with an inmate population. *Computers in Human Behavior*, 12, 159–166.
- Kveřton, P., Jeliřnek, M., Voborřil, D., Klimusova, H. (2007). Computer-based tests: the impact of test design and problem of equivalency. *Computers in Human Behavior* 23, 32–51.
- Lin C.Y., Xiao Z.W., Shen L., Zhang J.X., Weng X.C. (2007). Similar brain activation patterns for writing logographic and phonetic symbols in Chinese. *Neuroreport*. 2007 Oct 8 18(15), 1621-1625.
- Ministry of Education, Singapore. (2012) New iMTL Portal to Help Students Use MTL as a Living Language [Press release]. Retrieved from <http://www.moe.gov.sg/media/press/2012/08/new-imtl-portal-to-help-students.php>
- Terzis, V., & Economides, A.A. (2011). The acceptance and use of computer based assessment. *Computers & Education*, 56, 1032–1044.
- Xie, T. (2001). The digitalization of Chinese teaching and learning: The use of vocabulary processing software and its impact in teaching and learning. *Paper presented at International Conference on Internet Chinese Education 2001*, Taipei, Taiwan. In Chinese with title “手写 “还是”电写” —电脑输入中文引起的讨论. Retrieved from <http://www.csulb.edu/~txie/papers/E-writing.pdf>