An Experimental Study of Digital Reading Experiences

Huey-Wen Chou¹, Chien-Hong Chao²* and Chih-Hao Tu³

¹ Professor, Department of Information Management, National Central University, Taiwan
No. 300, Jhong-Da Rd., Jhong-Li District, Taoyuan City 32001, Taiwan

² Ph.D. Candidate, Department of Information Management, National Central University, Taiwan
No. 300, Jhong-Da Rd., Jhong-Li District, Taoyuan City 32001, Taiwan
*Corresponding author’s email: johnchao@iii.org.tw

³ Graduate Student, Department of Information Management, National Central University, Taiwan
No. 300, Jhong-Da Rd., Jhong-Li District, Taoyuan City 32001, Taiwan

ABSTRACT — With the popularity of the Internet and the development of information technology, digital reading has affected human reading styles. In essence, digital reading is different from conventional reading in many ways. The aim of this research focuses primarily on exploring the differences in reading behaviors among different digital reading devices. Results reveal that the reading experience on the Tablet PC is superior to that on the other two digital devices. Subjects in the Tablet PC group demonstrate the highest preference in terms of depth reading which implies that Tablet PC should be the most appropriate device for digital learning platform in the future. Discussion and suggestions are in the conclusions at the end of this paper.

Keywords — Digital reading, Reading behaviors, eBook reader, Tablet PC, Reading experiences.

1. INTRODUCTION

The rapid development of information technology (IT) and the Internet has enabled people to access a large amount of information in a short period of time. Reading has always been one of the most important ways for people to absorb information. In the past, the reading sources were papers and books. But, with the development of IT, today computers, e-book readers, tablets and mobile phones have evolved as the major reading carriers.

Finding Innovative New Digiservices (FIND) Institute for Information Industry) conducted a telephone interview to survey people aged 12 and over in Taiwan in 2010 (Tseng, 2010). In that study, they collected two thousand valid samples through sampling stratified by county, sex, and age. The results show that among digital reading devices, using a computer ranked the highest (50.5%), followed by mobile phone (21.1%), e-book reader, MP4/MP5 player, Personal Digital Assistant (PDA) and Portable Internet Device. Commonwealth Publishing Group conducted a Taiwan Digital Reading Survey in 2010 (Lin, 2010). The sample was people over 18 years old in Taiwan. It was found that nearly one-quarter of the respondents preferred to select e-books when e-books and paper books were both available.

The issues of “Digital Reading” behavior and attitudes have been widely discussed in literature. Huang (2010) explored the iPad reading experience with iPad users, and found they had a positive attitude towards digital reading. Another type of digital reading device is Kindle. Although Kindle is not as popular in the Taiwanese market as it is in the U.S. market, the development of this device is still indicative and worthy of attention and exploration. Hsiah (2010) pointed out that a small notebook computer with its size being similar to that of an e-book reader may perform the e-reader function. This study here then compares three types of devices – Kindle, iPad, and small notebook computers – to explore different digital reading behavior and experiences, hoping to better understand the current situation and reading experience of the digital reading market. In the past, for the study of digital reading, there was little discussion about the application of digital reading in deep reading. Therefore, this study used experimental design to explore and compare the feasibility and differences of digital reading devices in deep reading.

2. LITERATURE REVIEW

2.1 Shallow Reading and Deep Reading

The study classifies reading behavior into shallow reading and deep reading (Tuori, 1986; Liu, 2005; Wang & Wang, 2005; Chen, 2011). Shallow reading is also called rough reading or skimming; its purpose is to understand the information
with leisure and recreation. The process of shallow reading without thinking and with a jumping mode just pursues short-term visual pleasure and psychological pleasure. Therefore, fast-food, browsing, random, jumping, and fragmented readings are typical shallow readings. Shallow reading is replacing traditional reading, and it has become a new trend in reading. Liu (2005) questioned whether long-term shallow reading may reduce people's reading ability. In addition, shallow reading is also doubted for the purpose of leisure and entertainment, and has no significant help for learning (Wang & Wang, 2005).

Deep reading is also known as intensive reading. It refers to reading the content closely and treating reading as a level of learning and spiritual understanding. Through deep reading, the reader and the author establish a close tacit understanding and create dialogue in their two-way interaction. Chen (2011) defined deep reading as an intensive reading of articles that were selected through browsing and screening. Four kinds of deep reading are proposed – reading aloud by reading the contents of the articles word by word, reading and memorizing the contents of the articles, reading and understanding the author's thoughts and expressing his thoughts in one's own words, and intensive reading comprehension for the essence of the article. Macedo-Rouet et al. (2009) pointed out that digital reading behavior is mostly shallow reading. For content that requires deep reading, readers tend to print it out on paper and read. It is easy to mark on the paper and it causes less visual fatigue to read paper for a long time.

This study defines shallow reading and deep reading through a comprehensive literature review. Shallow reading is a non-linear reading behavior without a clear goal (Tuori, 1986). It is a reading behavior that often occurs in online reading. Deep reading means that the learner peruses and repeats reading, which makes readers able to understand the text content and describe the article specifically. This study intends to conduct research by experimental method, through selected professional articles as reading content, to dig out whether the reading experience of subjects using different digital carriers for deep reading is different.

### 2.2 Digital Reading and Digital Reading Devices

Digital reading refers to reading through digital media, so reading digital contents through a computer or other electronic devices can be called digital reading (Zhou, 2008). Digital reading contents refer to all digital contents, including electronic books (e-books or ebooks) or digital files or documents stored on CDs or on the Internet. The most commonly used digital devices include personal computers, notebook computers, e-readers, tablets, and smart phones. An e-book reader is a device that displays digital contents or e-book (Gibson & Gibb, 2011), which is represented in this study by the Kindle made by Amazon. The Kindle was launched with electronic ink (e-ink) as display technology in November 2007. E-ink is a type of electronic display for text that is intended to imitate the appearance of printed ink, typically used in e-readers. Past studies have shown that the reading device displayed by e-ink can effectively reduce the visual fatigue problem of users when reading for a long time (Siegenthaler et al., 2010). Amazon is the key player in many countries' e-book market (Gloria, 2018). The tablet PC is represented by Apple's iPad. According to ABI Research, in the third quarter of 2010, the global tablet sales volume was about 4.5 million units, – and the iPad accounted for 93% (4.2 million units). According to IDC (International Data Corporation) statistics, the global market sales of iPad in 2012 was 65.7 million units, with a market share of 50.2%, ranking first in tablet PCs (Kyle, 2013). Recently, Apple has tried to increase the visibility of its software so that its iPad and even the iPhone can gain more consumers' attention and earn hardware profits (Gloria, 2018). The above market data show that the Kindle and iPad are the mainstream in the digital content reading devices market.

A small notebook computer is a low-cost, small-sized, and streamlined computer, also known as Netbook. In 2008, ASUS and Intel jointly launched the series Netbook EeePC, which opened up the small notebook market. Later, various computer manufacturers, such as Acer, Dell, and Hewlett-Packard, entered the small notebook market, with their low-cost and portable features. In 2008, shipments reached 13 million units, forming a market segment (Kuo, 2009). Various manufacturers use low-voltage CPUs (Central Processing Units) to save energy. The screen size is 8.9 inches or 10 inches. We choose Acer Aspire one ZG5 digital device in this study.

In summary, both the e-book readers and the tablet PCs are very different from traditional digital reading devices (personal computers or laptops) in terms of operation, display technology, appearance and volume. This study employs these two digital reading devices and small laptops to understand the possible differences in the reading experience.

### 2.3 Reading Experience and Reading Comprehension

Reading experience refers to all the courses that readers actually read, including reading activities, explicit behaviors and intrinsic motivations. The main purpose of this study is to compare the subjects' digital reading experience on different reading devices. Therefore, the reading experience evaluation of digital reading is defined as the subject's workload level, usability assessment as well as the literal understanding, and comprehension of the reading content. The workload is the degree of mental working load perceived by the use of operating device when read, it is measured by NASA-TLX (Hart & Staveland, 1988). The NASA-TLX scale assesses mental demand, physical demand, temporal demand, effort, frustration level, and performance. The time period is not set, so the temporal demand is not included here. The remaining five question items are used to evaluate the working mental load level of using different digital devices to read. The device usability evaluation is based on the digital reading device evaluation project of Siegenthaler et al. (2010) to design the question item, including with respect to the overall design, browsing, convenience and
functionality of the device (both hardware and software functions), a total of 5 questions.

This study measures the extent to which the subject understands the content of the article after reading. Based on Pearson and Johnson’s study (1978), the reading understanding test was divided into two parts – literal understanding and content comprehension. The literal understanding means that the reader can recognize the Chinese characters of reading content while reading. Content understanding of the article means that the reader can understand the content of the article and has deeper thought with respect to the content in the article.

3. RESEARCH METHODOLOGY

In this study, laboratory experiments were conducted to design experimental tasks for graduate students, and repeated evaluation experiments were conducted to collect and analyze the subject's use behavior and reading experience on different devices. The deep reading content selected for the digital device in this study consisted of three articles in Chinese (Matthews, 2011 & 2012 [two]). This study used NASA-TLX to assess the degree of mental load. The assessment questionnaire was designed referring to Siegenthaler et al. (2010). For the content of each Chinese article, we also designed 5 reading comprehension questions to allow subjects to assess different devices after the experiment. By the students using devices and reading the article for evaluation and answering, as well as conducting in-depth interviews with the subjects after the end of the experiment, qualitative information was collected to understand the user's opinion about each digital reading device.

3.1 Experimental Design

The experimental samples used in the laboratory research were conveniently sampled, a total of 56 students studying for a master's degree at a university in the north of Taiwan. Before the experiment, the sample was pre-tested, and it was confirmed that the sample was subjected to digital reading experience to avoid the inability to operate the digital reading device due to insufficient reading experience. In addition, the study also excluded subjects with existing tablets or Kindle e-book readers by pre-testing to avoid the influence of past experience of the subjects. A total of 30 eligible subjects passed pre-test and participated in the experiment, with an average age of 24 years; participants were 20 men and 10 women.

Thirty subjects were grouped into 10 groups of A, B, ... , J in alphabetical order with each group having 3 subjects. The subjects were numbered according to the group order. For example, group A consisted of A1, A2, and A3. Each group of experiments took about 1 hour and 30 minutes. The experiment involved experimental instructions, practice using devices, reading experiments, questionnaires, and post-experimental interviews. The subjects used the order of the devices and the order of reading the articles in a Latin square design. Each subject used one of the different devices to read one of the selected three articles.

3.2 Experimental Process

The experiment consists of three units of article reading, and each unit is carried out for 20 minutes. Subjects read one article with different devices within three units. The first step during each unit takes 5 minutes to illustrate device use instructions and execute device trials. Subjects use the device to read the sample articles after the instructions, confirming that the subject can use the device smoothly and start reading with the device, and reading time is 15 minutes. When reading, participants are encouraged to use the various functions of the digital device to assist in reading, such as enlarging the article font, mark, note, etc. The subjects then complete the use of the experience questionnaire and the device function assessment scale and take a reading comprehension test. After completing the reading of three articles, the researchers conduct semi-structured recorded interviews with the subjects. This study designs an interview outline, based on the research questions and objectives, to collect qualitative data.

3.3 Device Specifications and Reading Software

The digital reading devices used in this experiment are Apple iPad 2 (iPad 2), Amazon Kindle DX (Kindle DX) and Acer Aspire one ZG5 (Aspire). Among three types of devices, Kindle DX uses built-in software and the other two devices use the Kindle Reading App. Through the same interface operating software, we focus on the nature of the device itself to reduce the impact on the reading experience caused by the functional differences of the software.

3.4 Reading Experience Assessment

The study used three scales to measure reading experience: Task Load Index (NASA-TLX), usability assessment, and a self-designed reading comprehension test. The five items on the NASA-TLX scale evaluated the mental load at the time of reading the article using different devices, measured by the Likert 7-point scale. The higher the score was, the greater the loads the subject was suffering when using the device. The device usability assessment was evaluated by the Siegenthaler et al. (2010) questionnaire to assess the device's overall design, browsing, hardware functionality, software functionality and convenience. A total of five questions was measured on a 5-point scale by Likert, "very much disagree"
(1 point) to “very much agree” (5 points). The higher the total score, the better the usability/availability of the device. The device availability scale and the mental workload scale items are shown in Table 1.

<table>
<thead>
<tr>
<th>Device Usability/Availability</th>
<th>Mental Workload Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>I like the design of this reading device.</td>
<td>I need a lot of thinking, remembering, searching it to read.</td>
</tr>
<tr>
<td>Reading with this reading device is easy and relaxing.</td>
<td>I need to spend a lot of time reading on the device.</td>
</tr>
<tr>
<td>The hardware features of this reading device are helpful to read.</td>
<td>I have to spend a lot of energy learning and operating this device.</td>
</tr>
<tr>
<td>The software features of this reading device are helpful to read.</td>
<td>I feel frustrated when I operate this digital reading device.</td>
</tr>
<tr>
<td>This reading device is quite easy to operate.</td>
<td>I am satisfied with the device’s performance in this reading experiment.</td>
</tr>
</tbody>
</table>

Referring to the research of Pearson and Johnson (1978), this study divides reading comprehension into two parts, literal comprehension and article content comprehension, and designs 5 questions for each article, 2 of which evaluate the reader’s literal understanding, and 3 questions evaluate the degree of understanding of the article content. Each question is single choice type from 5 choices. Since five domain experts reviewed and modified the questions, and then 30 graduate students were pretested, the questions should meet the surface validity and content validity.

The three readings used in the experiment were published in the BBC Knowledge magazine, the Small World Theory (Matthews, 2011), Gaia theory (Matthews, 2012), and the African Origin Theory (Matthews, 2012), which are categorized as science knowledge. After a pilot test, it is confirmed that the difficulty level of the article content is appropriate for the target subjects. The number of words in the three articles is 4923 words, 5019 words, 4656 words, respectively, and the maximum difference in word length is about 360 words, which can avoid the influence of the length of the article on the length of the experiment. Taking the article “Small World Theory” as an example, the evaluation of literal understanding level is “Who is the scholar that first studied the small world theory?” and the semantic understanding level is "What are the key factors for the formation of the small world effect?"

4. RESEARCH FINDINGS AND DISCUSSIONS

4.1 Results Analysis

The reading comprehension questions for each of the three reading articles are 5 questions with each question worth 20 points. The study found that the average score for reading comprehension using Kindle DX and Acer Aspire was both 64, and iPad 2 was 60.67, showing that the subjects’ reading comprehension score with the iPad 2 was slightly lower than with the other two devices. ANOVA results found that the three types of devices in the NASA-TLX scale have significant differences in physical demand, effort, frustration and performance. This scale shows that iPad 2 is superior to the other two devices, as shown in Table 2. Subjects using iPad 2 for digital reading paid lower mental and physical effort, but got lower reading comprehension score.

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Devices</th>
<th>Average score</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mental demand</td>
<td>iPad 2</td>
<td>4.27</td>
<td>Kindle DX</td>
</tr>
<tr>
<td>Physical demand</td>
<td>2.77</td>
<td>4.93</td>
<td>2.90</td>
</tr>
<tr>
<td>Effort</td>
<td>2.57</td>
<td>4.30</td>
<td>2.37</td>
</tr>
<tr>
<td>Frustration</td>
<td>2.33</td>
<td>3.93</td>
<td>2.50</td>
</tr>
<tr>
<td>Performance</td>
<td>3.23</td>
<td>4.07</td>
<td>3.63</td>
</tr>
</tbody>
</table>

*p<.5, **p<.01, ***p<.001

The scores of the three devices in the usability assessment differ significantly, as shown in Table 3. The iPad 2 is significantly better than the other two devices in each aspect of the usability evaluation. It shows that the touch-screen operation mode is intuitive, convenient, and easy to read. It does help digital reading. The size and weight of the iPad 2 makes it easy to read in a variety of different positions.
4.2 Interview Results Analysis

After completing the reading experiment, a one-to-one, semi-structured interview with the subjects was conducted. The interview content was recorded. The interview results of the 30 subjects were summarized as follows.

(1) Purpose and content of reading: Sixteen subjects believed that their reading behavior was related to the purpose in reading and the reading content. If the purpose of reading is for deep learning and taking notes is needed, then paper-based reading is preferred; otherwise, reading from digital devices is preferred.

(2) Reading occasions: Fourteen subjects preferred digital reading in a mobile location or outdoors because digital reading is more convenient than paper reading.

(3) Reading device preferences: Twenty subjects indicated that iPad 2 reading brought the best digital reading experience, because the convenience in operation, the quick response time, and the touch screen were beneficial for annotation. Ten subjects chose the Aspire’s small notebook with familiar interface, including mouse and keyboard. No subjects raised the merits of the Kindle. The details are explained based on the operating experience, device response time and visual burden points as follows.

A. Operational experience: Twenty-nine subjects thought that reading with Kindle DX was difficult because the keyboard operation of the device was complicated and cumbersome, and this had an impact on reading speed, reading comprehension and reading interest. The experience of using Kindle was not good.

B. Response time: Twenty-five subjects indicated that Kindle DX was slow to respond and it caused interruption of reading and affected reading interest.

C. Visual load: Past studies have shown that e-ink can effectively reduce the visual burden. In our experiment, 17 subjects thought that the display mode would affect the reading behavior, but the other 13 thought that the display mode would not influence their reading behavior.

(4) Choice of future devices: Twenty-four subjects chose a tablet, and six subjects chose a notebook.

4.3 Summary

The results of the experimental data analysis showed that the participants used the annotation function during the reading process. The number of annotations ranged from 1 to 10, indicating that not many participants had habits of using the annotation function. From the NASA-TLX assessment results and subsequent qualitative interviews, the characteristics of the three devices were analyzed as follows.

(1) iPad 2: Both NASA-TLX scales and device usability evaluations are superior to those measures for the other two devices. The burden of digital reading using the iPad 2 is low and gives the subject a better reading experience. Most subjects believe that the touch operation mode is intuitive and easy to read. It can be inferred that operating in touch mode is helpful for digital reading behavior. In addition, some subjects believe that the size and weight of the iPad 2 make it easy to read in a variety of different postures, which helps one to read, so the size and weight of the device itself also affect the reading behavior.

(2) Kindle DX: The NASA-TLX scale and device usability evaluation are lower than for the other two devices, but the average understanding score is higher than the iPad 2, and the same as the Aspire. The reason for the lower average reading experience score from qualitative interviews is that most subjects think that Kindle DX has a slower response, which leads to interruption of reading and operation through the keyboard on the device, which is cumbersome and difficult to use.

(3) Aspire: The operation method is the same as that of the notebook computer, only the size and performance of the display screen is slightly lower than that of the general notebook computer. On the NASA-TLX scale, there is no significant difference between the Aspire and the iPad 2, but the device overall design score in the device usability assessment is lower than that of the iPad 2. The post-interview indicated that the subjects’ reading experience in Aspire was better than with the iPad 2, which may be because the subjects are familiar with the keyboard and mouse operation mode. Therefore, the user’s familiarity with the device operation is significantly affected in the user’s digital reading behavior. But in the choice of future devices, only 6 subjects think that the notebook is a better choice for digital reading, so the familiarity with the device itself is not the main consideration with respect to which device to choose.
A comparison among the reading comprehension tests, the NASA-TLX scale and the usability evaluation of the three types of digital devices find that the touch operation mode of the iPad 2 can effectively reduce the user's mental workload, but the reading comprehension score is slightly lower than with use of the other devices. The Kindle DX scored lower on the usability assessment score, and the NASA-TLX scale was higher (i.e., it requires greater effort), but the reading comprehension score was slightly higher than the iPad 2. As for Aspire, the users are familiar with its operation mode. The evaluation of the above items is between the iPad 2 and the Kindle. Therefore, this study infers that the user's familiarity with the operation mode of the device and its convenience and ease of use will affect the mental workload and experience of using the device.

5. CONCLUSIONS

5.1 Research Conclusions

This study finds that iPad 2 provides a better reading experience in this experiment. The main reason is that the operation mode of the touch screen is simple and convenient, and it can be operated and read smoothly without much effort. Especially in annotation and text input, the control method gives the user access to the experience of reading on paper, which is also beneficial for reading. In addition, the reaction speed of the reading device is also one of the factors affecting the reading experience. Slow reaction speed will interrupt reading, which will increase the user's reading burden. Therefore, if the e-book reader can improve the reaction speed and operation mode, and match the advantages of the e-ink for reading, it should create a unique and superior reading experience.

In reading comprehension, the experiment found that there was no significant difference among the three devices, but the Kindle DX and Aspire scored better than the iPad 2. Although the iPad 2 is better at assessing mental workload and availability, it does not respond to reading comprehension, and the reason is worthy of further study.

In the experiment, it is recommended that the subjects can use various auxiliary reading functions of the device, such as magnifying the font, marking, and annotating. However, this study does not monitor the experimental process, and it is impossible to discuss the relationship between these functions and learning experience. Future research on device-assisted reading can provide insight into the relationship between device function and reading experiences.

5.2 Research Limitations

Although the research process and research design in this study are complete, they are limited with respect to time, manpower and material resources. There are the following research limitations. First, because graduate students are selected as a convenient sample to conduct experiments in the implementation of reading experiments, the experimental samples’ ages, education levels and other characteristics unable to be randomly selected and lack representativeness. Therefore, the sample diversity is insufficient and the inference of the study results is limited. Secondly, due to the innovation of the software and hardware functions of the reading device, it takes time for the research input and the experimental analysis. Thus, this study is based on the most common mainstream hardware and software specifications at that time. It is recommended that follow-up researchers can continue to study the effects of the software and hardware specifications of the latest features on the digital reading devices.

5.3 Suggestions for Future Research

Today's digital reading has gradually matured both in the development of devices and in the market. This study explores the application of emerging devices in recent years, such as iPad 2, Kindle DX in digital reading, and designs experiments with difficult content to estimate the feasibility of using digital reading for deep reading. Suggestions are made for the promotion of digital reading in education and learning. The following research directions are in the future research.

The laboratory experiment of this study selected users with digital reading experiences to read three science articles with similar lengths and similar themes on three digital reading devices. This study adopted Latin square experimental design to evaluate the reading understanding level and reading experiences. Because of limited time, manpower and resources, it was not possible for users to read for a longer period of time. This study only chose factors of reading comprehension, physical demand, mental load, and device usability assessment to explore their possible impacts on different reading devices. There are many more other factors remain to be further examined.

In addition, this study found that the differences in reading comprehension scores among different devices are small. Therefore, it is suggested that future research can carry out deep reading experiments on different devices and for a longer period of time to explore whether there will be reading comprehension differences in deep reading among different devices. It is not convincing to infer that which digital device is more suitable for deep reading in terms of device usability and mental demand scale. In the future, if the deep reading experiment can be carried out for a long time, the digital reading device with e-ink display screen may provide a more comfortable reading environment and conditions than the backlight tablet PC, which is worthy of further discussion by subsequent researchers. According to the results of this experiment, the tablet computer found in this study enables a better reading experience in digital reading. Whether it can be inferred to have better reading comprehension for deep reading for a long time needs more follow-up studies.
6. ACKNOWLEDGEMENT

This study is supported in part by the National Science Council of the Republic of China under contract number NSC 101-2511-S-008-007-MY2. We would like to thank all the participants in the questionnaire and the experimental interviewees for providing valuable data in this study.

7. REFERENCES