Evaluation of e-Readers: A Preliminary Analysis

Ujala Rampaul and Tom Gedeon

Research School of Computer Science
Australian National University
{u5051329 | tom.gedeon}@anu.edu.au

Abstract. Evaluation of new consumer goods, particularly electronic, are often done from the perspective of analysing the latest models, comparing their advantages and disadvantages with respect to price. This style of evaluation is often performed by one of a few product experts on a wide range of features that may not be applicable to each user. This study instead used a scenario-based approach to evaluate a number of e-readers, mimics a user interested in a new product or technology with a limited budget. The objective is to evaluate from a purchasing perspective the quality and usability of e-readers available within that budget range. The e-readers evaluated were operated in multiple ways, which suggests that the interface design should cater for users with different levels of experience with technology. The results of a large user study with over 70 participants, shows that the popular brands do not necessarily produce the best products in terms of e-readers. We found that navigation within books to be the most significant differentiator between the eReaders in our scenario based evaluation process.

1 Introduction

In 2007, Amazon launched the Kindle in the United States and it sold out in under 6 hours (Patel, 2007). Prior to the Kindle launch there was not a huge demand for e-readers as earlier releases were the Rocket eBook in 1998, the Sony Libriè in 2004 and Sony Reader in 2006. These earlier devices were unsuccessful, perhaps because they were expensive, had technological limitations and lacked available content. The Kindle was different because it was affordable, consumers could purchase a wide variety of e-books via the wireless data connection and it was comparatively easy to use.

E-readers with e-ink screens simulate the experience of reading a paper book and are not multifunctional devices. The advantage of an e-reader is that it is easier to read even in direct sunlight, does not consume a lot of battery, is lighter in weight, permits undisturbed reading of an e-book and the eyes do not fatigue compared to devices with backlit LCD displays.

In the United States, the cost of e-readers has continued to fall, making them more accessible to purchase. Data from Pew Research Centre over a three-year period from May 2010 to January 2014 indicate e-reader ownership has grown over 20%, see Figure 1 on ownership growth of e-readers and tablets in the United States. Overall, 50% of Americans now have a dedicated handheld device – either a tablet computer like an iPad, or an e-reader such as a Kindle or Nook for reading e-content, and $^1/_3$ of US adults now own an eReader (Zickuhr and Rainie, 2014).

Figure 1: eReader ownership in the US

The motivation of this project was to evaluate from a purchasing perspective the quality of e-readers that could be purchased within a limited budget. Shoppers of e-readers and e-reading devices are often influenced by advertising, reviews conducted by computer and consumer sites such as PC Authority and Choice, as well as opinions of friends and families. The commercial reviews compare the performance, features, file formats and specifications of the latest e-reader models against the market price to identify the best value for money.

To obtain e-readers for this study a budget based on the cost of a new paper book ($25) was established and multiplied to set the budget range of approximately $50-75 (excluding shipping costs) for the purchase of an e-reader. Establishing a budget reduced the range of e-readers that could be obtained i.e. either, earlier, e-reader models with a higher specification or later models with a relatively low specification. We argue that if we focused on the quality of the devices available within this budget range, then the differences in the e-reader model and specifications were not limiting factors. The experiment requires that a realistic range of e-readers of a high quality could be obtained from eBay and other auction sites for an effective comparison.

The following nine e-readers were selected and mostly purchased from within Australia except for the Sony and jetBook, at a lower price range than originally estimated being $35 -$70. This includes a selection of e-readers having buttons only, touch screen only or a combination of both.

e-reader	Cost in AUD$	Study device label
Amazon Kindle 4G	57.85	C
Barnes & Noble Nook	69.90	A
Ectaco jetBook	61.28	G
Elonex 621EB	41.00	F
EZReader EZ601	55.00	I
iRiver Story HD	52.00	E
Kobo Touch	60.00	B
Pandigital Novel	36.85	H
Sony PRS-600	44.88	D

Table 1. Cost of second-hand e-readers

2 Background

The rise of e-books and e-readers are part of a larger story about the shift from printed to digital material. Early research focused on academic contributions towards education where students used e-readers in the classrooms to read textbooks.

There have been numerous approaches taken when evaluating both ebooks and ebook readers in projects such as Superbook and Electronic Books On-screen interface (EBONI). (Gibson & Gibb, 2011).

Other studies included university and public libraries investigating strategies to support the borrowing of e-books and e-readers. "In particular, many academic libraries since begun pilot projects using a variety of different reader devices to investigate the possibilities for simplifying and innovating the content and related services they offer to users in light of the technologies that are available. Countless libraries have experimented with offering lending programs for their devices. Many libraries have undertaken these initiatives on their own whereas others have partnered with either Sony electronics or Amazon to design and carry out their projects" (Behler & Lush, 2010).

To establish design guidelines for e-readers (Pearson, Buchanan, & Thimbleby, 2010) conducted a study to determine the usability concerns with user interface of the e-readers using human computer interaction (HCI) principles. The experiment was carried out using the Kindle 2, Sony PRS 600 and Sony PRS 300. These three e-readers had similar screen type (e-Ink technology), screen size (six inches) and resolution (600 x 800). The Sony PRS 600 was a combination of touch and button device whereas the Kindle 2 and Sony PRS 300 were button devices.

Based on the guidelines and principles for ergonomics, consistency, completeness, page numbering, book marks, annotation and magnification, the Kindle was found to be better than Sony PRS-300 and PRS-600. The positioning of the buttons in the Kindle made it easier to turn pages for both left and right-handed users. The Kindle had a full QWERTY keyboard which supported web browsing and location identifiers in relation to the file instead of page numbers. The availability of full annotation and zooming was only available on the Sony PRS-600. User feedback on the three e-readers was leveraged from Amazon's online customer review and weighted on a Likert scale rating. This could have attributed for the higher ratings for the Kindle, as it is produced by Amazon.

The study by Gibson and Gibb (2011) evaluated four second-generation e-reading devices, namely Sony PSRS505; Bookeen Cybook 3g; iRex iLiad and Asus Eee PC 105 HA, with 33 participants. Their evaluation measured weight, quality of display, size and robustness of the devices based on a five point Likert scale rating.

On overall impressions and functionality, the Asus Eee netbook was the best, followed by Sony PRS 505, iLiad and lastly Cybook. The Asus Eee netbook was the preferred choice because participants were familiar with the design and layout compared to the dedicated e-readers (Gibson & Gibb, 2011).

Both the Sony PRS505 and Booken Cybook were similar in screen size (six inches) and resolution (600 x 800 pixels) whereas the iRex iLiad had a larger screen size (eight inches) and a larger resolution (768 x 1024 pixels). All three devices had screens with e-Ink technology. The Asus Eee PC 105 HA, a netbook had a LCD screen, a much larger screen size (ten inches) as well as a large resolution (600x1024 pixels) and weighed seven times more than the Booken Cybook, the lightest (175g) device in the trial. This meant the refresh rate for turning the page were faster in the netbook compared to the e-readers using e-Ink technology. This was an unequal comparison between three dedicated e-readers and a netbook.

Usability was favourable towards the e-readers for lightness and portability, readability and ease of use. Participants commented, "the screen was not wearing on the eyes" (iLiad) and that it was a "straightforward operation" (Sony), and that "the non-glare screen made the text as easy to read as ink" (iLiad) (Gibson & Gibb, 2011).

Based on reported sales and market share in the United States, Richardson & Mahmood (2012) studied five leading e-readers namely Kindle 3G; iPad 1G; Nook; Kobo N647; and Sony PRS950. Their objective was to identify the advantages and disadvantages of e-readers and "to compare and contrast the most popular (i.e. bestselling) devices against a comprehensive, if not exhaustive, set of technical specifications as well as qualitative judgment of users" (Richardson & Mahmood, 2012).

The results from Richardson and Mahmood's study showed that the Kindle was the most popular device, though some participants commented on poor navigation. Some 47% of the participants that undertook this study owned a Kindle, accounting for a probably biased popularity of the Kindle. There was no evidence for the technical comparison nor did they elaborate what it consisted of. The comparison between the iPad and the e-readers was probably unjust as the tablet is a multi-function device, whereas e-readers are designed to read and purchase e-books.

The NextMedia eReading Project conducted a qualitative and quantitative study on reading behaviours and patterns on e-readers including the effects of age groups. The objective of the main project was to encourage Finnish consumers to adopt and grow an e-reading community. This eReading Project consisted of multiple smaller projects that focused on different aspect of the study such as usability of e-reading devices, types of content, emotional and cognitive reactions.

Heikkilä (2011) usability study of e-reading devices was based on a scenario, which involved the participants recording their experiences of opening and using a newly purchased e-reading device.

A number of pages from a paper book was the benchmark for their evaluations. The same content being available on the e-reading devices. Over a week, seventeen participants recorded their reading times, places they had read and their experiences on using the e-readers. The time-based assessment tasks included opening the device; finding a book; finding a specific spot in the book and changing the font size. Heikkilä used the following e-readers in his study Kindle (included both Kindle 2 and Kindle 3); Sony PRS600; iPad 1G; BeBook Neo; Booken Cybook Opus; Elonex eBook and Samsung galaxy.

The outcome of the usability study was the creation of a conceptual model of an e-reader labelled Acola. The Acola was a combination of touch sensitive and gesture savvy pad on an e-ink device. The pad would handle page turning and skimming. Swiping to the left would turn the page forward, swiping to the right would move one page backwards. Swiping fast two times successively, would turn two pages, three times: three pages etc. Swiping with two fingers simultaneously would move you between chapters. A menu and an OK button could be situated in the upper and lower border of the pad (Heikkilä, 2011).

There is some similarity between Heikkilä's (2011) study and this study in the use of a scenario for the evaluation of the e-reading devices. It is interesting that his study evaluated two tablets and four e-readers to create their ideal e-reader.

Siegenthaler, et al., (2012) conducted a study to with the following reading devices Sony PRS505, Sony PRS600 and an Apple iPad, to determine effects of touch screen technology. Twelve participants tested the three devices sequentially within a session based on a set of time related tasks such as open a book, open a specific page within the book, highlight a sentence, find a highlighted sentence, delete a highlighted sentence, change the orientation of the layout and increase the font size. The participants rated the above tasks on the navigation, design, handiness and handling based on a Likert scale rating.

The results showed "that e-reading devices with touch screens correlate with better navigation rating" and touch screen technology also has its advantages in terms of a higher intuitiveness and flexibility for adaptations of the navigation (e.g., due to firmware updates) compared to devices with static buttons. (Siegenthaler, Bochud, Wurtz, Schmid, & Bergamin, 2012). The evaluation of handiness and handling was ambiguous, for example the related questions were "How handy do you rate the reading device?" and "how easy was it for you to handle the reading device?" There was no clarification to what this meant and the evaluation questions were vague.

We earlier identified the differences between an e-reader, which is solely for reading e-books whereas a tablet is a multifunction device making this evaluation an unequal comparison.

Siegenthaler, et al, (2012) argued, "E-ink technology has low power consumption, thereby increasing battery life and allowing for a more lightweight device. Another advantage is that e-ink devices can be used outside without glare being a big issue. However, e-ink screens have some disadvantages, most of them are black and white and the pages do not refresh as quickly as devices with an LCD screen".

The disadvantage of tablets is that they cost considerably more than e-readers, the LCD screens are susceptible to glare, they do not provide a comfortable reading experience, are heavier and the battery does not last as long.

E-ink technology is designed to emulate printed books. Independent studies on the reading behaviours measured by eye tracking found no difference between a paper book and a dedicated e-reader (Nielsen, Kindle 2 Usability Review, 2009) and (Siegenthaler, Wurtz, & Groner, Improving the usability of e-book readers, 2010).

Similarly, analysis of the reading behaviours between a dedicated e-reader and an e-reading device (tablet) found no significant difference as measured by eye speed (Nielsen, iPad and Kindle Reading Speeds, 2010) and (Siegenthaler, Wyss, Schmid, & Wurtz, 2012). However, the iPad "actually scores slightly higher in user satisfaction" (Nielsen, iPad and Kindle Reading Speeds, 2010).

Please see Figure 2 which maps the reading devices that are common to this study and to other studies identified. Most studies shared one common e-reader, for example Richard & Mahmood

(2012) share the Nook, Siegenthaler, et al.,(2012) shared the Sony PRS-600, Siegenthaler, et al., (2010) share the jetBook and Pearson et al., (2010) share the Sony PRS-600. The exception is Heikkilä (2011) where there two common e-readers namely the Sony PRS-600 and the Elonex.

Please see Figure 3 which maps the tasks that are common to this study and to other studies identified. All studies share at least one common task, with Siegenthaler et al, (2010) and (2012) evaluating two similar tasks namely, open a document (open a book) and increase font size. The other exception is Heikkilä (2011), where four similar tasks were evaluated that is turn the power on (open device), navigate to first document (find a particular book), increase font size (change font size) and navigate to specific section within the book (find a specific place within the book). All studies have commented on readability.

No previous studies have been conducted to evaluate from a purchasing perspective the quality and usability of e-readers purchased within a budget of up-to $70. This study proposes to conduct the evaluation of the e-readers based on a scenario with two experiments.

1. Evaluation of nine e-readers that are button interface devices only, touch screen only and a combination of both to determine which device is the best
2. Evaluation of e-readers with touch screens

3 Methodology

The Human Research Ethics Committee at the Australian National University (ANU) approved this study. Participants from the ANU College of Engineering and Computer Science were recruited for the experiments.

The participants evaluated the devices in pairs so one could act as the scribe and observe, while the other operated the device. On completing the evaluation of the first device, the participants swapped roles to evaluate the next device. This grouping encouraged discussion amongst themselves on their observations and interactions with the e-readers. In total, 72 senior students (3rd year HCI course participants, Honours, Masters and PhD) took part in the evaluations. Please see Table 1 for the list of 9 devices evaluated. Each device was allocated an alphabetical letter based on the arrival sequence for easy identification in the randomisation matrix.

The randomisation was similar to a Latin square where two sets of the e-readers were labelled from A to I and arranged in such a way that no row contained the same letter twice. The pairs were balanced, so that the number of times a letter came first or second was the same. There were 36 pairs of e-readers for 72 participants to evaluate, which meant each device was evaluated four times as a first device and then four times as a second device. Table 2 shows an example of the device 'A' evaluated eight times with the other devices.

BA	BH	HE	IF
GE	**CA**	BI	GC
DC	IE	**DA**	BG
DI	EC	CH	**EA**
AF	HD	CF	HF
CB	**AG**	GD	CI
GF	DB	**AH**	ED
FD	IH	EB	**AI**
IG	FE	HG	FB

Table 2. Matrix of the e-reader randomisation pair for device A.

The evaluation commenced when each participant pair were given the first e-reader from a randomised pair, and a questionnaire. The participants received an overview of the study and its

objectives. After performing the tasks on the first e-reader, the tasks were repeated on the second e-reader.

The questionnaire included a scenario, the evaluation tasks based on Likert-like numerical scale, open-ended questions to gain an insight on improving the device and comments after each device evaluated. It also gathered optional personal details including the name, gender, age, year of study, familiarity with e-readers and follow-up contact details for purposes of clarification or removal from the experiment in adherence with The Human Research Ethics Committee guidelines of the Australian National University. Lastly, the questionnaire requested the subjective opinions and reasons on whether one of the two devices was preferred.

The evaluations were designed around the same scenario to provide context to the experiment. The scenario was based on a user finding a gift received a couple months ago. Included in the box was a note to register the device prior to use. The participant had to register the device without the assistance of the user manual, which was based on a series of tasks, to be accomplished with the assistance of a fellow student, who had similarly received an ereader as a gift. For both experiment, the participants had to assume the battery on the e-readers were fully charged, as it was not a requirement of the scenario.

The tasks was measured on a Likert-like summative scale rating ranging from the most positive to the least positive:

- 5 = Very good
- 4 = Good
- 3 = Ok
- 2 = Bad
- 1 = Very bad

The tasks were

1. To turn the device on
2. Navigate to a document (called "Somedevice XY User Guide")
3. Open the document
4. Increase the font size
5. Navigate to specific section to find the 'model number'
6. Navigate to a second document ("Dates-2013" and use the model number).
7. Assess readability on the screen

Explicit instructions were given to not use the *recent history*, *continue reading* or the '*date*' search in the e-readers. Most of the e-readers provided such functionality, which manufacturers seem to consider of such high utility that in general it is not possible to turn off these modes, which makes a repeated experiment such as ours impossible without this instruction to not use such features.

Additional subjective open-ended questions were requested, on likes and dislikes about the e-readers; previous use or ownership of an e-reader and the model and if they were regular users of a smartphone; tablet or laptop. These questions were to gain a better understanding into the user's experience with mobile devices.

4 Results

Statistical analysis was performed using F-statistics based on a repeated measures ANOVA for the within factor of e-readers. A critical $p < 0.05$ was used for statistical significance in all analyses. Each of the following sections report the ANOVA results for that question based on the usability tasks measured on a Likert-like numerical scale rating from 1 (very bad) to 5 (very good) on the e-readers.
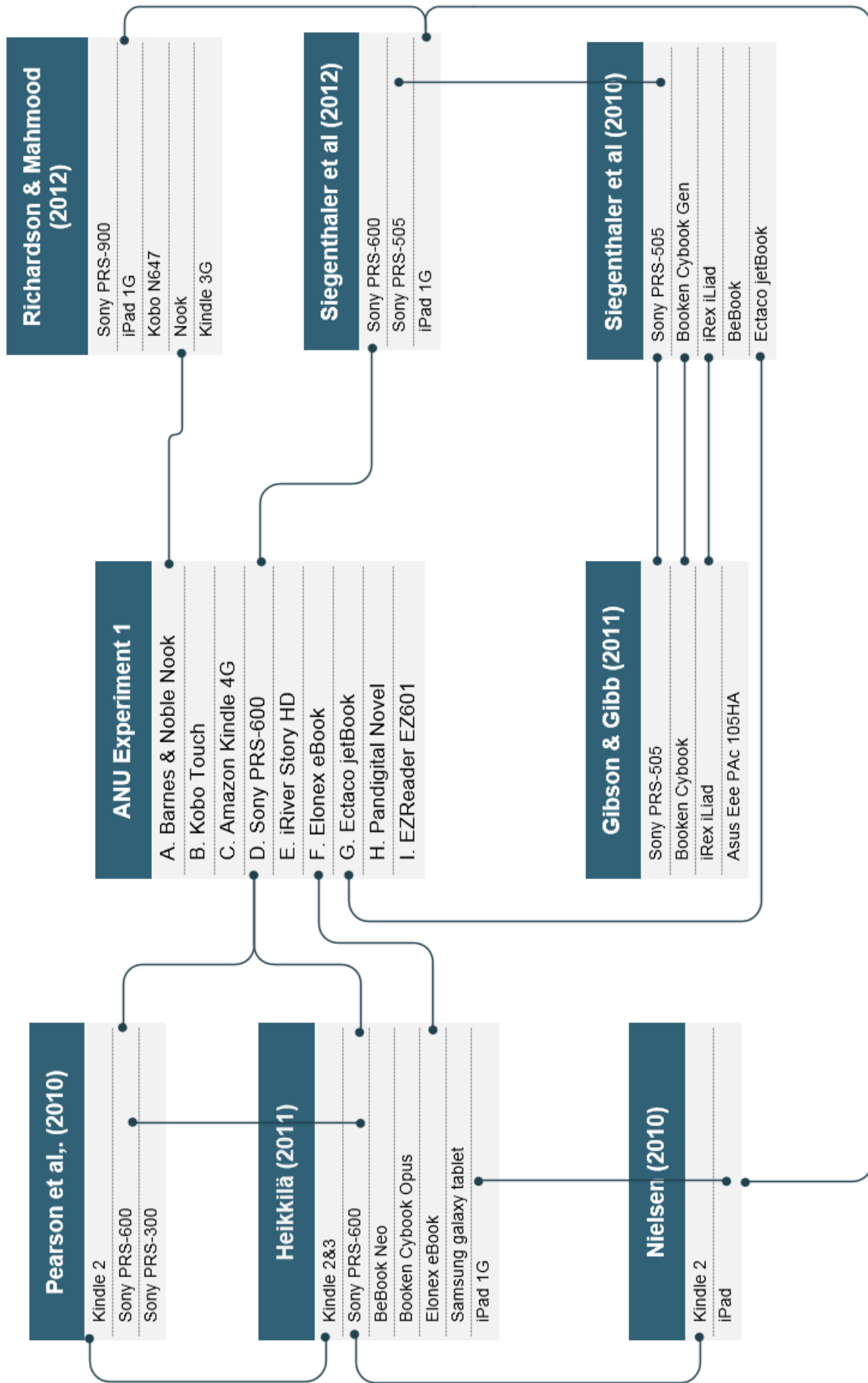
Figure 1. Map of common reading devices between this study and previous study
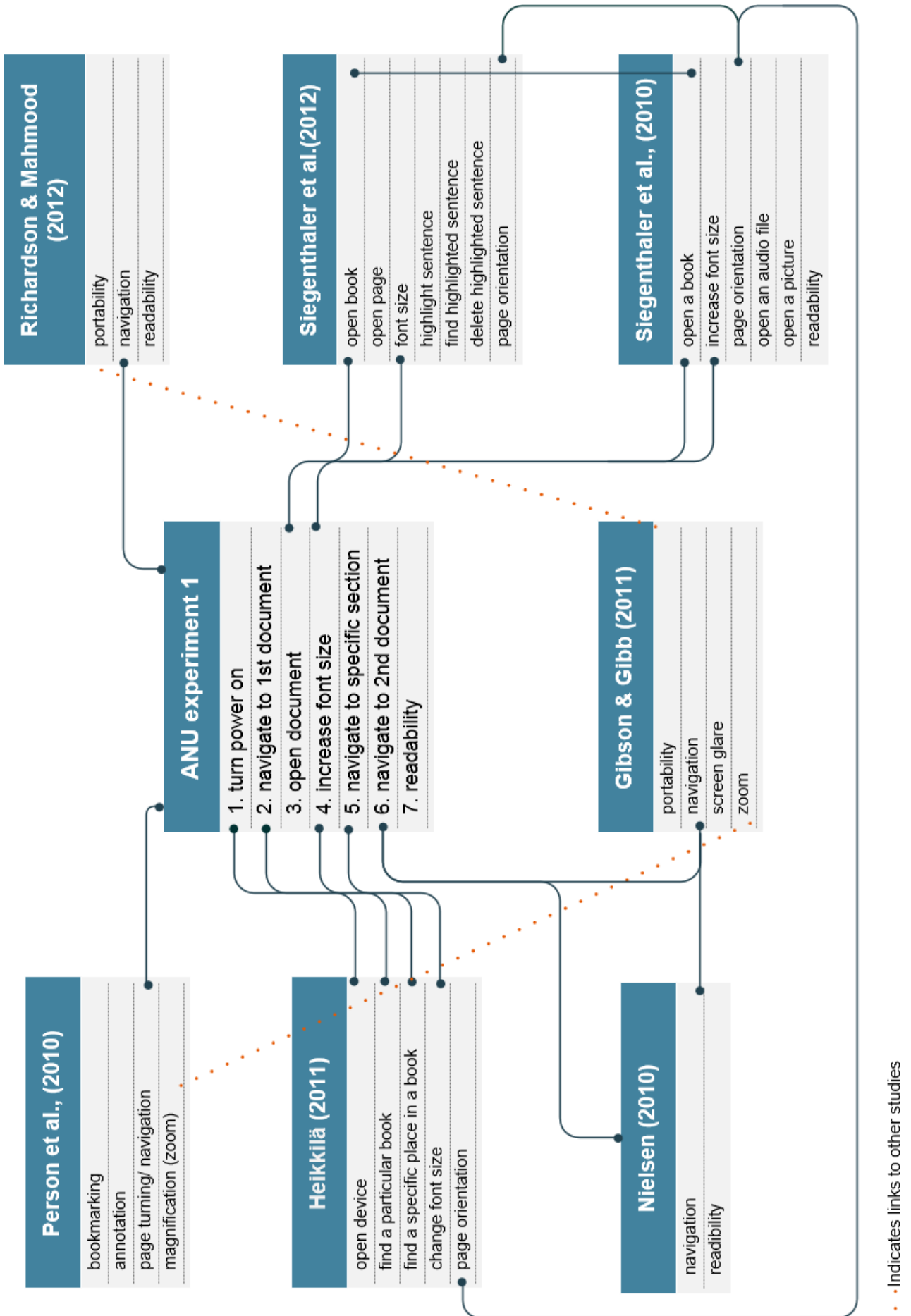
Figure 2. Map of tasks between this study and previous studies (all comment on readability)

4.1   Turning on the e-reader

There were no significant differences found in the responses to the question on powering on the e-readers. Table 3 shows the results and ranks for the e-readers. The jetBook, Pandigital and the Kobo are ranked at the top. There are four e-readers (i.e. Nook, iRiver, Elonex and EZReader) ranked at the middle in equal fourth. The Sony and the Kindle were ranked at the bottom, however there were no significant differences found in the e-readers for turning them on.

e-Reader	Mean	SD	Rank
jetBook	4.1	0.83	1
Pandigital	4.0	0.76	2
Kobo	3.9	1.17	3
Nook	3.8	1.28	=4
iRiver	3.8	1.28	=4
Elonex	3.8	0.71	=4
EZReader	3.8	1.28	=4
Sony	3.4	1.06	8
Kindle	3.1	0.83	9

Table 3. Mean Ranking and Standard Deviation (SD) for turning the e-readers on

4.2   Finding the first document

There were no significant differences found in in response to the question on finding the first document. Table 4 shows the results and ranks for the e-readers. Sony and Kobo ranked in the top for finding a document whereas the iRiver, Pandigital, jetBook, and Nook were ranked in the middle. EZReader and Kindle were ranked at the bottom but there was no significant difference found in all e-readers. Note that each eReader was pre-loaded with over 70 books downloaded from the Project Gutenberg (2013) website of books now in the public domain. The selection was eclectic, consisting of popular fiction (e.g. Burroughs (1917)), mythology (Werner, 1922), craft (de Dillmont, 1886), history (Gibbon, 1776-1789) and classical philosophy (Plato, 350 BC).

e-Reader	Mean	SD	Rank
Sony	4.0	0.53	1
Kobo	3.5	1.41	2
iRiver	3.3	1.16	=3
Pandigital	3.3	1.49	=3
jetBook	3.0	1.31	5
Nook	2.9	1.46	6
Elonex	2.5	0.76	7
EZReader	2.4	1.60	8
Kindle	2.3	0.89	9

Table 4. Mean Ranking and Standard Deviation (SD) for finding the document

4.3 Opening the document

An analysis of variance showed that the responses to the question on opening the document was significantly different between the e-readers, $F(8,63) = 3.21$, $p = 0.004$. Table 5 shows the results and rankings for opening the document. The Kobo, Kindle and the iRiver were ranked at the top. The Sony and Elonex were ranked equal fourth. The jetBook and Pandigital ranked equal sixth. The EZReader and Nook rank at the bottom. This shows there is a significant difference in the e-readers to perform the task of opening a document.

e-Reader	Mean	SD	Rank
Kobo	4.6	0.52	1
Kindle	4.5	0.93	=2
iRiver	4.5	0.76	=2
Sony	4.4	0.74	=4
Elonex	4.4	0.74	=4
jetBook	4.3	1.39	=6
Pandigital	4.3	0.89	=6
EZReader	3.5	1.20	8
Nook	2.8	1.16	9

Table 5. Mean Ranking and Standard Deviation (SD) for opening the document: **significant**

4.4 Increasing the font size

An analysis of variance showed that increasing the font size was significantly different between the e-readers, $F(8,63) = 2.25$, $p = 0.03$. Table 6 shows the results and rankings on the e-readers for increasing the font size. The Pandigital, jetBook and iRiver were ranked at the top three positions. The Kindle and Elonex were ranked equal fourth. The Nook was ranked at sixth position. The Kobo and EZReader were ranked at the bottom two positions as equal eighth. This shows a significant difference was found in the e-readers to perform the task of increasing the font size.

e-Reader	Mean	SD	Rank
Pandigital	4.6	0.52	1
jetBook	4.1	1.13	2
iRiver	4.0	1.07	3
Kindle	3.9	1.46	=4
Elonex	3.9	1.13	=4
Nook	3.4	1.06	6
Sony	3.3	0.71	7
Kobo	2.8	1.75	=8
EZReader	2.8	1.49	=8

Table 6. Mean Ranking and Standard Deviation (SD) for increasing font size: **significant**

4.5 Navigate to specific section for find the 'model number'

An analysis of variance showed that navigating to a specific section within the document was significantly different between the devices, $F(8,63) = 3.43$, $p =0.002$. Table 7 shows the results and rankings on the e-readers for navigating to a specific section within the document. The jetBook, Pandigital and the Kindle ranked at the top. The Sony, iRiver and the Kobo ranked in the middle and the Elonex, Nook and EZReader ranked in the bottom positions. This shows a significant difference in the e-readers for navigating to specific section within a document.

e-Reader	Mean	SD	Rank
jetBook	4.0	0.93	1
Pandigital	3.6	1.19	2
Kindle	3.5	1.07	3
Sony	3.4	1.41	4
iRiver	3.3	1.28	5
Kobo	2.8	1.39	6
Elonex	2.6	1.30	7
Nook	1.9	1.46	8
EZReader	1.6	0.92	9

Table 7. Mean Ranking and Standard Deviation (SD) for navigation to section 3: **significant**

4.6 Navigate to a second document

There was no significant differences found in navigating to a second document with the e-readers. Table 8 shows the results and ranks for navigating to a second document. The iRiver, jetBook and the Sony ranked in the top three positions. The Pandigital, Elonex and the Kobo ranked at the middle and the Nook, Kindle and EZReader ranked at the bottom three positions. Navigating to a second document was not significantly different in the e-readers.

e-Reader	Mean	SD	Rank
iRiver	4.0	0.93	1
jetBook	3.9	1.81	2
Sony	3.8	1.04	3
Pandigital	3.6	0.92	4
Elonex	3.4	1.69	5
Kobo	3.1	1.64	6
Nook	2.9	1.46	7
Kindle	2.6	1.41	8
EZReader	2.5	1.31	9

Table 8. Mean Ranking and Standard Deviation (SD) for navigation to second document

4.7 Readability on the screen

There were no significant differences found in response on assessing the readability. Table 9 shows the results and ranks for readability. The iRiver was ranked the highest, followed by Kobo, Kindle

and jetBook in second position and the Pandigital in the fifth. The Nook and EZReader were both ranked equal sixth. The Elonex and Sony ranked lower but no significant difference were found between the e-readers for readability.

e-Reader	Mean	SD	Rank
iRiver	4.5	0.76	1
Kobo	4.0	0.53	=2
Kindle	4.0	0.93	=2
Jetbook	4.0	0.93	2
Pandigital	3.9	0.35	5
Nook	3.8	0.71	=6
EZReader	3.8	1.28	=6
Elonex	3.5	1.20	8
Sony	3.3	1.28	9

Table 9. Mean Ranking and Standard Deviation (SD) for readability

4.7 E-Reader preference

For the question, "Did you like one of the two devices better?" an analysis of variance showed that the responses were significantly different between the e-readers, $F(8,54) = 2.33$, $p = 0.03$. Table 10 shows the results and rankings of the participants' preference when comparing e-readers. The Kindle and iRiver ranked equal first, followed by Sony, jetBook and Pandigital in equal third position. The Elonex was ranked sixth, Kobo seventh and both Nook and EZReader ranked at the bottom. This indicates a significant difference in the e-readers that participants' preferred the most.

e-Reader	Mean	SD	Rank
Kindle	0.6	0.52	=1
iRiver	0.6	0.52	=1
Sony	0.5	0.53	=3
jetBook	0.5	0.53	=3
Pandigital	0.5	0.53	=3
Elonex	0.4	0.52	6
Kobo	0.3	0.46	7
Nook	0.0	0.00	=8
EZReader	0.0	0.00	=8

Table 10. Mean Ranking and Standard Deviation (SD) for e-reader preference: **significant**

4.7 Pairwise comparison on e-readers with significant differences

Pairwise comparison is commonly used to estimate preference values of finite alternatives with respect to a given criterion. This is part of the model structure of the analytical hierarchy process, a widely used multi-criteria decision-making methodology. The main difficulty is to reconcile the

inevitable inconsistency of the pairwise comparison matrix elicited from the decision makers in real-world applications (Choo & William, 2004)

The elimination of designs or candidates can change the tabulated rankings of those designs or candidates that remain under consideration. The determination of which design is "best" or which candidate is "preferred most" may well be sensitive to the set of designs considered (Dym, Wood, & Scott, 2002).

The overall ranking we used based on the qualitative and quantitative mean rankings of the questions about e-readers that indicated a significant difference as shown in 4.3 opening a document, 4.4 increasing the font size, 4.5 navigating to specific section within the document and 4.8 e-reader preference. Table 11 shows the overall ranking of e-readers where significant difference was noted.

e-Reader	Qualitative ranking	Quantitative ranking	Overall ranking
Kindle	=1	3	=1
Pandigital	=3	1	=1
iRiver	=1	4	=3
jetBook	=3	2	=3
Sony	=3	5	5
Elonex	6	6	6
Kobo	7	7	7
Nook	=8	8	8
EZReader	=8	9	9

Table 11. Qualitative, quantitative and overall ranking

The qualitative rankings were based on the number of *yes* responses to the question on whether one of the two devices was better. The quantitative rankings were based on Likert-like numerical scale rating on the individual tasks and participants' responses in questions where a statistically significant difference was found in the responses.

The qualitative and quantitative rankings are similar, within one rank for the majority of e-readers. This suggests that our rankings can be relied on, as the two forms of data are largely consistent. The exceptions are the Kindle (1th in qualitative ranking and 3rd in quantitative ranking); Pandigital (3rd in qualitative ranking and 1st in quantitative ranking); iRiver (1st in qualitative ranking and 4th in quantitative ranking) and Sony (3rd in qualitative ranking and 5th in quantitative ranking). These differences are discussed below.

5   Discussion

Interaction with e-readers should be simple, effective and a pleasant experience as it is emulating a paper book.

The power ON/OFF buttons on the Sony and the Kindle was not easily noticeable. On the Sony, three similar buttons of the same shape and size were located at the top edge of the device and the label was not distinct due to the reflective surface and the angle at which the device was commonly held. A comment on the Sony was "doesn't stand out from the other buttons on the same area". The power switch on the Kindle was positioned on the lower edge of the device and the button was not labelled. Some of the comments on the Kindle power button were "button hard to find at the bottom - unusual position"; "hard to find switch button"; and "hard to locate the button, button un-labeled". The ANOVA results showed no significant differences in the e-readers in response to the question for turning the e-readers on.

Some devices such as Sony, Elonex, jetBook, iRiver, Pandigital and EZReader stored the recently read e-books in the current reading or recent history folder. This extremely useful feature allows the reader to go straight to their e-book without being distracted with cumbersome

navigation and searching. In the experiments there was only one of each e-reader, when an e-reader was available for the next pair to evaluate there was not sufficient time to refresh the reading history folder, as it generally requires a full factory reset. It was a compromise to leave the documents within the folder especially if alternate methods for finding the document were supported. The recent history was not an obvious feature to all participants. Users were instructed that if the e-reader 'accidentally' started in the target document, then they should back out of it and navigate to it.

Typing a search query in the Kindle was cumbersome and required too much effort to complete the full query as it involved navigating the keyboard via the 5-way controller. Prevention of keyboard slips and incorrect spelling was not supported and frustrating to the user. An autocomplete feature would be a suitable design solution to combat the complexity in typing and error management.

Buttons are designed to afford pressing to trigger a visible change on screen. The search button on the EZReader was unresponsive and disconcerting as the users were unsure how to continue. If a button is displayed it must be associated to an action otherwise it leads to user frustration and a waste of valuable real estate on the interface. Although the participants encountered some difficulty with some of the e-readers on finding the document, the ANOVA results show there was no significant difference in the e-readers.

Opening a document was found to be either intuitive on some e-readers or challenging on other e-readers. This is also reflected in the ANOVA results, which showed a significant difference in the e-readers to support opening a document. For example, tapping on the document title opened the document on the Kobo, Sony and Pandigital. The iRiver had two "enter" buttons, one situated above the keyboard and the other within the keyboard for easy selection. The center buttons on the multi-controllers of the Kindle, Elonex and the jetBook behaved as the "enter" key.

Comments on the EZReader for opening a document were "hard to find", "tried the arrows keys first", and "a bit slow". The "enter" button on the Nook touch panel was not identifiable because it did not meet a user's concept of a button and it was not labeled. A comment on the Nook for opening the document was "open button is not obvious, task a bit to find. Fact screen is not touch is not obvious". The "enter" key could not be distinguished if it was a radio button or a decorative element.

The ANOVA results for increasing the font size showed a significant difference in the e-readers to accomplish this task. The Pandigital supported alternate methods to change the font size for example, within an opened document from previous task, zoom icons (both plus and minus) were available on the bottom right corner of the screen and selecting the menu button displayed a toolbar with additional access buttons such as "*My Library*; *Dictionary*; *TOC*; *Bookmarks*; *Go To*; *Font size* and *Next*" on the top of the screen.

The font size adjust button was visible on the multi-controller of the jetBook, the keyboard on the iRiver and a stepper on the Elonex. The option to change the font size in the Kindle was accessible via the "menu" button.

Comments on the Nook on changing the font size were "menu a bit tricky to scroll through" and "not intuitive to find". The Kobo has a multi-touch screen which meant each area of the screen was reserved for a function, for example tapping in the middle or top right of the screen turned a page forward and tapping on the top left of the screen turned the page backwards. A tap to the bottom of the screen was essential to unfold the font size icon button and a comment was "hidden toolbar, obvious icon".

In the EZReader a "font" option was displayed via the "option" button but this was ambiguous as only the font type could be changed and not the font size. Comments on changing the font size on the EZReader were "no idea where to navigate"; "had to select font first, that wasn't any good so we kept looking through other menus" and "can't find the setting". Each e-reader had its own methodology to accomplish the task of changing the font size which accounts for the difference identified in the analysis of variance calculation.

In the Kobo the "table of contents' options was available via the concealed toolbar, similar to accessing the font size option. Comments on the Elonex were "first tried to use 'explore' button and press keyword 'model number', no response came out. Easy to find when using 'content' button". The 'content' button was accessible via the 'menu' button.

There were no visual cues for a search or table of contents feature in the Nook and user's found it to confusing. Comments on navigation in the Nook were "couldn't figure out how to scroll the screen", "not found" and kept pressing the wrong button (left button sent pages forward)".

Similar to the Nook there were no visual cues in the EZReader to gain access to the specific section within the document. Comments on the EZReader to navigate to the section within the document were "cannot find it", "we cannot find it", "needed 15% hint for go to page", "it's impossible to find that paragraph" and "can't find the document". These variations for navigating in the e-readers to a specific section within a document were also identified in the analysis variance results.

Navigating and opening a second document reinforced the learning experience from the previous task for finding and opening a document and no significant differences were found in the ANOVA results.

There was no significant difference for readability. Almost all the e-readers had e-Ink screen types. Pandigital had SiPiX and jetBook had a LCD screen type, but these did not affect the readability. SiPiX technology uses white particles suspended in black ink while e-Ink technology has both black and white particles suspended in a clear fluid. A comment on Sony was "screen is dull and not enough backlight for contrast" but it was not significantly different compared to the other e-readers.

A significant difference in the analysis of variance results were found when asked which e-reader was preferred from the pair evaluated. This was then used to conduct a comparison to determine which e-readers was overall significantly better, based on counting the number of times each device was considered better than the other device.

There were four e-readers for which their qualitative rankings were more than one-step different in the quantitative rankings. The Kindle, Sony and the iRiver were all higher by two to three steps and the Pandigital was two steps lower. This could be due to the Pandigital being an unknown brand compared to the Kindle and Sony as popular brands, though this does not explain the iRiver which showed the highest difference between the qualitative (1st) and quantitative (4th). It is possible that the brand recognition is the primary answer for the Kindle and Sony. For the iRiver, it appears that the overall interface is visually appealing to the users to account for the high qualitative rankings but the functional aspects and the aesthetics of the interface are not in complete harmony to meet the user's expectation.

Comments on the iRiver and EZReader were that the keyboards were redundant; the buttons were too small and overall not very useful. A full list of participants' comments is available on request. Here is a summary of the comments:

1) The e-readers were either non-intuitive or intuitive
2) Some e-readers were found to be user friendly and easy to use
3) Expectations were for faster and more responsive feedback
4) E-readers with un-labelled buttons and use of icons were confusing
5) The lack of contrast colour of buttons and interface was deceptive
6) Preference for a device with touch screen interface

A "grounded theory" approach was considered to interpret the qualitative data but the comments were found to be too short and insufficient to code and to form concepts of the code, for a meaningful analysis. Grounded theory is a systematic generation of theory from systematic research (Corbin & Strauss, 2008).

In this study, three e-readers, the Pandigital, the Sony and the Kobo used Touch technology. The Pandigital and the Sony are also button-operated interfaces, while the Kobo was touch only. The

qualitative rankings for these devices were Pandigital (3rd) and Sony (3rd) and quantitative rankings were Pandigital (1st) and Sony (5th). The Kobo is consistent with both its qualitative and quantitative ranking at 7th. It is clear from these tasks highlighted in 4.9 pairwise comparison, that button operated interfaces are consistently better, that is the three devices which had touch interfaces are distributed among the top, middle and bottom. Further, the only full touch e-reader (Kobo) is among the bottom three.

It is worth noting that most new devices available on the market in 2015 are mostly touch devices, and over 60% of the participants commented that they either preferred a device with touch screen interface, or would improve the device by adding a touch screen interface.

In experiment 2, the preliminary analysis of variance calculation showed a significant difference in the e-readers only for navigating and opening a second document. This could be attributed to the instruction on to not use the 'date' search. That is, mentioning 'search' could have predisposed the participants to use the search feature to find the documents. The second document could be found only by the title search. Content search was not possible because the second document was a very short document, with very little content to search on.

The search and table of contents features has not been investigated in previous studies, most likely because it is not a feature associated with paper books. However, on an electronic device a search feature is critical to finding information quickly. The search feature should be investigated in the future.

The overall ranking is based on the qualitative and quantitative mean rankings of the e-readers with a significant difference from the experiment 1 (Q3-Q5) and experiment 2 (Q6) to understand the difference between the sets.

6 Conclusion

The ANOVA results showed significant differences between the e-readers for the tasks of opening the document, increasing the font size, navigating to a specific section within the document, and user preference on the e-reader from the device pair evaluated.

In this study the tasks was measured on a Likert-like numerical scale rating the e-readers which was used to calculated the Mean, Standard Deviation and Analysis of Variance (ANOVA). From ANOVA results, a further comparison on the qualitative and quantitative rankings were used to determine which e-readers were overall better.

In the overall ranking the Kindle, Pandigital, iRiver and jetBook were in the top four positions. The Kindle, iRiver and jetBook are button-operated e-readers whereas the Pandigital is combination of buttons and touch. E-Readers with buttons are intuitive to use. Users know when they have selected a physical button whereas tapping a button on a touchscreen is not always instantaneous due to the speed of response and the user may continue to tap the screen.

This study concluded an analysis on nine e-readers. The next stage of this project is to extend the evaluation for touch based e-readers as they seem to be dominating the current market for e-readers, requiring more data collection and further analysis. Exploring the direction for further studies could involve:

- Updating the scenario to evaluate tasks for continuous reading
- Evaluating the e-readers by concealing the brand names
- Evaluating the usability of an e-reader to read one-handed whilst drinking, coffee, holding an umbrella, standing on a bus or train and so on
- Evaluating the search and table of contents features in e-readers
- Investigate different styles of documents (e.g. Rho & Gedeon, 2000) and readability
- Directly investigate reading behaviour with e-readers using eye gaze (e.g. Vo et al, 2010)

References

Anon., 2013. A bright future. Choice, June, pp. 54-57.

Behler, A. & Lush, B., 2010. Are You Ready for E-readers?. The Reference Librarian, 52(1-2), pp. 75-87.

Burroughs, E.R. (1917): A Princess of Mars, A. C. McClurg U.S., xii, 326 pp Choo, E. U. & William, W. C., 2004. A common framework for deriving preference values from pairwise comparison matrices. Computers & Operators, 31(6), pp. 893-908.

Corbin, J. & Strauss, A., 2008. Basics of Qulaitative Research. 3rd ed. s.l.:Sage.

de Dillmont, T. (1886): Encyclopedia of Needlework, original title: Encyclopédie des ouvrages des dames.

Dym, C. L., Wood, W. H. & Scott, M. J., 2002. Rank Odering Engineering Designs: Pairwise Comparison Charts and Borda Counts. Research in Engineering Design, Volume 13, pp. 236-242.

eInk (2014): E Ink: Technology: How E Ink Works, Online: eink.com/how_e_ink_works.html, [Accessed 1 Sep 2014]

Gibbon, E. (1776-1789): The History of the Decline and Fall of the Roman Empire, volumes 1-6, Strahan & Cadell, London.

Gibson, C. & Gibb, F., 2011. An evaulation of second generation ebook readers. 29(3), pp. 303-319.

Heikkilä, H., 2011. eReading User Experience: eBook devices, reading software & contents. eReading Media Use, Experience & Adoption, 10 January, Volume 1.

Nielsen, J., 2009. Kindle 2 Usability Review. Online: www.nngroup.com/articles/kindle-2-usability-review/ [Accessed 10 March 2014].

Nielsen, J., 2010. iPad and Kindle Reading Speeds. Online: www.nngroup.com/ articles/ipad-and-kindle-reading-speeds/ [Accessed 10 March 2014].

Patel, N., 2007. Kindle sells out in 5.5 hours. Online: www.engadget.com/2007/11/21/kindle-sells-out-in-two-days/ [Accessed 5 April 2014].

Pearson, J., Buchanan, G. & Thimbleby, H., 2010. HCL Design Principles for eReaders. Swansea and London, ACM, New York, USA, pp. 15-24.

Plato (circa 350 BC): Euthyphro, publisher unknown.

Project Gutenberg (2013): Free ebooks, Online: gutenberg.org [Accessed 1 October 2013].

Rainie, L. & Smith, A., 2013. Tablet and E-reader Ownership Update. Online: www.pewinternet. org/2013/10/18/tablet-and-e-reader-ownership-update/ [Accessed 8 April 2014].

Rho, Y.J. and Gedeon, T.D., 2000. Academic articles on the web: reading patterns and formats. *International Journal of Human-Computer Interaction,12*(2), pp.219-240.

Richardson, J. & Mahmood, K., 2012. eBook readers: user satisfaction and usability issues. Emerald insight, 30(1), pp. 170-185.

Siegenthaler, E. et al., 2012. The effects of touch screen technology on the usability of e-reading devices. Journal of Usability Studies, May, 7(3), pp. 94-104.

Siegenthaler, E., Wurtz, P. & Groner, R., 2010. Improving the usability of e-book readers. Journal of Usability Studies, November, 6(1), pp. 25-38.

Siegenthaler, E., Wyss, M., Schmid, L. & Wurtz, P., 2012. LCD vs E-ink: An analysis of Reading Behaviour. Journal of Eye Movement Research, pp. 1-7.

Vo, T., Mendis, B. S. U., & Gedeon, T., 2010. Gaze pattern and reading comprehension. In *International Conference on Neural Information Processing* (pp. 124-131). Springer Berlin Heidelberg.

Werner, E. T. C. (1922): Myths & Legends of China. New York: George G. Harrap & Co. Ltd.

Zickuhr, K. & Rainie, L., 2014. E-Reading Rises as Device Ownership Jumps. Online: www. pewinternet.org/2014/01/16/e-reading-rises-as-device-ownership-jumps/ [Access 5 04 2014].