Multi-Level Storytelling App

Emma Dickson
062442792

August 2016

Computer Science
School of Mathematical and Computer Sciences

Dissertation submitted as part of the requirements for the award of the degree of MSc in Creative Software Systems
Abstract:

Reading for leisure has been proven by numerous research projects to be one of the best ways to improve literacy levels, yet time spent reading is in decline throughout the UK population. Worryingly, this corresponds to a drop in literacy levels, with 16% of the adult population barely able to read a simple text.

So, how do we get people to read more? This project consists of the design, development and evaluation of an interactive multimedia eBook app, aimed at measuring whether such apps can increase reading motivation and enjoyment amongst certain areas of the population. One of the key features of this app will be provision for variable difficulty levels to accommodate less able readers. Through user-centred design, with an iterative three-stage development and evaluation process, the aim is to create an eBook app that is engaging and motivating for even poor readers.
Author's Statement

I, Emma Dickson, hereby confirm that this work submitted for assessment is my own and is expressed in my own words. Any uses made within it of the works of other authors in any form (e.g., ideas, equations, figures, text, tables, programs) are properly acknowledged at any point of their use. A list of the references employed is included at the conclusion of this report.

Signed:

Emma Dickson

18/08/16
Acknowledgements

This project would not have been possible without the support of several people and groups, whose contributions I would like to acknowledge here.

Firstly I would like to thank my project supervisor, Prof. Rob Pooley, for his regular support and advice.

Secondly mention must be given to the staff and pupils of Echline Primary School, whom I was privileged to visit to conduct user testing sessions. In particular I would like to thank head teacher Dorothy Crawford, class teachers Sarah Dickson, Fhiona Emslie, Ingrid Ingram and Jennifer Miller, and all of the children from P3/4, P4, P5 and P5/6 who took part.

Thanks also go to Adrian Hurt for allowing me the loan of all of MACS’s Android tablets, without which the school testing session would not have been possible.

I would also like to thank everyone who took part in my initial consultation and Experiment 2 evaluations for giving their time to provide feedback and suggestions.

Finally thanks go to my partner Alistair and my family for their support throughout this project.
Introduction

The aim of this project is to create an app which can be used to determine whether a multi-level multimedia reading platform can be of use in encouraging struggling or reluctant readers to read more often.

The following sections will:

- Explore the existing research literature related to literacy, reading habits and interactive storytelling
- Set out the project’s aims and hypotheses
- Detail the project plan and schedule
- Examine the app development process
- Present the design and results of two user testing experiments
- Discuss any problems and limitations to the project
- Present overall results and conclusions
- Suggest further work
Contents

Abstract .................................................................................................................. 1
Author’s Statement ................................................................................................. 2
Acknowledgements ............................................................................................... 3
Introduction .......................................................................................................... 4
Contents ................................................................................................................. 5
1 – Literature Review ............................................................................................. 7
  1.1 – Introduction ................................................................................................. 7
  1.2 – “Background Reading” ................................................................................. 8
  1.3 – What Motivates People to Read, and What Books Do They Enjoy? .......... 11
  1.4 – How do eBooks relate to reading motivation? .......................................... 17
  1.5 – What About Interactive eBooks? ................................................................. 20
  1.6 – Interactive Storytelling in Research ............................................................. 22
  1.7 – Summary - The Story from Here... ................................................................. 28
2 – Requirements Analysis ..................................................................................... 29
  2.1 – Project Brief & Mission Statement ............................................................... 29
  2.2 – Stakeholder Requirements ......................................................................... 30
  2.3 - Software Requirements .............................................................................. 31
3 – Methodology ................................................................................................... 35
4 – Project Plan ....................................................................................................... 37
5 – App Development ............................................................................................. 41
6 – Experiment 1 ..................................................................................................... 45
  6.1 – Implementation ............................................................................................ 45
  6.2 – Results – Likert Data .................................................................................. 46
  6.3 – Results – Comment Feedback ..................................................................... 55
7 – Experiment 2 ..................................................................................................... 63
  7.1 – Implementation ............................................................................................ 63
  7.2 – Results ......................................................................................................... 64
8 – Discussion ......................................................................................................... 68
  8.1 – Project Implementation .............................................................................. 68
  8.2 – Questions and Answers ............................................................................... 69
  8.3 – Results Summary ......................................................................................... 71
8.4 – Future Work ........................................................................................................ 71
9 – Conclusions ........................................................................................................ 73
10 – References ......................................................................................................... 74
1 – Literature Review

1.1 – Introduction

This literature review will examine several cultural and technical factors motivating and influencing the development of this interactive eBook project. These will include the importance of reading as a core skill, the link between regular leisure reading and literacy level, and how motivating people to read more can therefore have a positive effect at both an individual and societal level. This review will also consider the current state of eBook reading, in terms of prevalence, reader habits and the relative success of eBook adoption across different usages. Finally, existing research into the field of interactive storytelling and its relevance to the project will be analysed.
1.2 — “Background Reading”

Reading is considered to be an essential life skill, yet 16.4% (OECD, 2012) of the UK adult population – equating to around 5.2 million people in England – can be described as “functionally literate.” In real terms, this means that their reading abilities are limited to simple, short texts, or the reading level expected of an average 11-year-old (National Literacy Trust, 2016). This is by no means a localised problem either: of the 23 G20 countries surveyed in the OECD Skills Outlook report (OECD, 2013), an average of 12.2% of the population were ranked in the same category (1 or below). Overall scores by country are shown below in Figure [1]:

This lack of reading skills can have a serious effect on daily life, with low literacy linked to below-average rates in areas ranging from earning potential to political engagement and even health and wellbeing (OECD, 2012), (Gleed, 2013). Those with limited literacy are also

<table>
<thead>
<tr>
<th>Country</th>
<th>% Below Level 1</th>
<th>% Level 1</th>
<th>% Level 2</th>
<th>% Level 3</th>
<th>% Level 4+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Japan</td>
<td>1.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finland</td>
<td>6.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Netherlands</td>
<td>2.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sweden</td>
<td>6.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norway</td>
<td>2.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estonia</td>
<td>6.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slovak Republic</td>
<td>6.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flanders (Belgium)</td>
<td>5.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canada</td>
<td>6.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Czech Republic</td>
<td>6.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>1.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denmark</td>
<td>6.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Korea</td>
<td>6.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>England/N. Ireland (UK)</td>
<td>1.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>1.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>United States</td>
<td>4.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Austria</td>
<td>1.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poland</td>
<td>6.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ireland</td>
<td>6.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>6.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cyprus</td>
<td>17.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spain</td>
<td>6.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>6.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1 - Reading Proficiency in Surveyed Countries by OECD Proficiency Score Band (OECD, 2013, Fig. 2.1, p.63)
effectively deprived of access to a wide range of both informative and cultural resources.

As those who have trouble reading are less likely to spend time reading for leisure, a vicious cycle becomes apparent. As listed below, studies have shown that from an early age, those who read often for their own enjoyment have far higher mean literacy scores than their peers, whilst struggling readers may become locked into the habit of non-reading due to negative experiences. As mentioned in (Clark and Rumbold, 2006), the reciprocal relationship between reading often and improving as a reader is well established (Cunningham and Stanovich, 1998), and holds just as true with current figures.

In his analysis of reading-proficiency research in US elementary schools, (Allington, 2014) stresses the role of the volume of voluntary reading done by children as one of the most significant influences on developing literacy. This link was supported across several of the studies analysed, in which time spent on reading practice, as opposed to exercises in individual word recognition or phoneme decoding, was found to be the only factor tested to significantly improve subjects’ reading capabilities.

A 2006 meta-analysis of research on reading for leisure conducted for the National Literacy Trust in the UK (Clark and Rumbold, 2006) concurs, with numerous studies mentioned concluding that reading for pleasure has a positive relationship with:

- Reading and writing test results
- Text comprehension and grammar skills
- Wider vocabulary
- Favourable attitudes towards reading
- Confidence in reading abilities
- Enjoyment of reading in later life
- General knowledge
- Understanding of other cultures

This pattern also holds with older readers; whilst the majority of the studies considered dealt with younger children, McGeown et al. (2015) focused their study on high-school age students (11-16). They examined the relationship between subjects’ level of motivation and their reading behaviours and several variables benchmarking reading ability. They found that in general, highly-motivated readers and fiction readers performed significantly better at reading comprehension, summarisation and text reading speed, after correction for baseline word reading ability. Pupils’ confidence in their reading ability was more closely linked to reading skill, whilst reading value (perception of reading as enjoyable, useful and important) related more closely to levels of engagement. This suggests that with the right support and attitude, a poor reader could be just as engaged with reading materials as an advanced one.
The same holds even for post-school-age adults: a study of low-literacy adult learners in (Rodrigo, Greenberg and Segal, 2014) showed that students enrolled on a literacy skills programme improved their reading habits (reading more frequently and sticking to a book) and motivation levels, in turn promoting more frequent reading practice and literacy skill gain.

However, volume of reading alone is only part of the story. To gain the best results in advancing reading skills, both the variety and difficulty levels of the material available must be considered. The findings of McGeown at al. (2015) state that reading of longer texts, especially fiction books, were the most consistent in predicting reading skill levels in teenagers, in comparison to other text sources. Topping et al. conducted a large-scale survey of school assessment data for US Grade 1-12 readers (Topping, Samuels and Paul, 2007), and found that combining difficulty level of texts read and volume of reading had the most significant positive effect on literacy level improvement, particularly in older readers. From this, it may be inferred that an eBook fiction app which provides several difficulty levels, allowing more competent readers to challenge themselves whilst still providing a variety of content at a simpler level, would be of benefit to readers through its ability to satisfy both requirements.

Another interesting statistic arising from this research is that Topping et al. found that readers’ initial literacy levels were less of a factor in the improvement measured across the assessment period than the students’ classroom placements, indicating that proper support and guidance in reading practice may outweigh a lower initial reading level. This fits with discussion by Allington (2014) that teaching strategies chosen for supporting struggling readers, in addition to the number and variety of books available to learners, can have a significant influence on learning outcomes. Providing an additional resource to aid in reading practice can therefore be seen as another weapon in the educator’s arsenal.
1.3 – What Motivates People to Read, and What Books Do They Enjoy?

With a causal link between regular reading and literacy level firmly established across existing research, the question which inevitably arises is: how do we get people to read more?

Numerous studies have collected data on reading habits and motivation; the following examines a cross-section of these representing different age groups in order to build a picture of just what motivates people to read for fun, what they enjoy reading and what might prevent them reading.

Starting with school-age children, one resource of use is the Scholastic Kids and Family Reading Report, which is based on data from an annual YouGov survey in the US, quizzing parents and children on kids’ reading habits and motivations (Scholastic Inc. and YouGov, 2015). This is a large-scale survey (over 2500 responses) with the respondent group selected to be demographically representative of the US population. Children’s responses were also checked against parent involvement in helping them answer to correct for any parent influence. This study therefore provides one good baseline data set for comparison.

Regression analysis of their survey data showed that, in order of significance, the biggest motivational factors which influenced children aged 6-17 who were frequent readers were:

- Enjoyment (at least an 8/10 rating given for reading enjoyment)
- Having parents who were frequent readers
- Holding the belief that reading for fun is important

For younger children (ages 6-11) reading aloud at home and having been read aloud to by parents at a young age were also found to be influential motivators. In teens (12-17) additional factors of significance were: independent reading time during school, parents who encouraged them, having a large home library, and knowing their reading levels.

Frequent readers across the whole 6-17 age group also reported less time spent engaging in screen-based activities such as watching TV, video games and social networking, though any causal direction in this link is unclear. A majority of subjects in the 6-17 group also reported a decline in their enjoyment of reading as they got older, with the most common reason cited for this being that there were more activities competing for their time.

There was also an overwhelming preference (averaging 91% across age groups) for reading books that the child has chosen themselves. With 73% of children surveyed saying they would read more frequently if they could find more books they enjoyed, these preferences should be taken into account when designing reading resources.

When asked what they want in a book, young readers overall cited funny books as the biggest appeal, followed by “lets me use my imagination” and “tells a made-up story” (Scholastic Inc. and YouGov, 2015). The latter results fit with the findings of (Dungworth et
al., 2004), in which the authors found in their study of Year 5 (8-9-year-old) pupils that the most popular reasons given for reading were emotional: firstly a sense of relaxation and enjoyment, and secondly a feeling of being “transported into the story.”

When broken down by age, humour remains the favourite book genre according to (Scholastic Inc. and YouGov, 2015), but there was also an upward trend with age in appreciation for books which featured romance, related to the reader’s own life experiences and gave a feeling of escape from reality.

In (Nestlé Family Monitor, 2003), a large-scale weighted survey of 11-18-year-olds carried out in England and Wales, humour was still popular, but it was narrowly beaten to the top spot by fantasy fiction. This was perhaps due to the popularity of the Harry Potter and Lord of the Rings series at that time, due to the publicity surrounding the respective movie releases.

Teens were also found to feel more motivated overall towards fiction reading than factual, and preferred both over assigned school reading (McGeown et al., 2015), reinforcing the idea that free choice of reading materials is a significant factor in motivation levels.

Factual learning was not entirely in disfavour, however, with a 2003 study of 11-18-year-olds, when asked why they read, responding:

“55% of these young people stated that books help them understand different people/cultures, 40% wanted to learn more about new subjects, and 33% stated that books encouraged them to try new hobbies.” (Nestlé Family Monitor, 2003, in Clark and Rumbold, 2006, p14).

The subjects of this study also cited enjoyment and relaxation as major reasons, similar to the findings of (Dungworth et al., 2004), in younger students.

Moving onto older readers, a 2014 study (Huang et al., 2014) examined the reading habits of US college students (subject ages not supplied, therefore estimating most college students to be 20-24 based on the largest college student age group by census figures (U.S. Census Bureau, 2014)). The research showed that students spent an average of 7.72 hours on assigned class reading, 4.24 hours on leisure reading, and 8.95 hours on the internet, including social media, blogs and email. Internet usage time was measured as a whole, not confined to text-based activities, so the amount of online time spent engaging with text in some form is difficult to estimate.

Unsurprisingly, online reading was given as subjects’ preferred reading format, followed by magazines and news, graphic novels and traditional novels. This contrasts sharply with younger readers’ preference for fiction, but may be understandable as many of the students said they had little free time for reading, perhaps leading them to favour shorter texts. The
overall picture given by the data collected is one of a strong preference for web-based, social and informative reading material amongst this age group.

This study may not be taken as a totally representative sample of young adults, however; as the authors highlight, the demographic is limited, consisting of largely white students from one school of one college campus, and the number of participants, particularly in the qualitative side of the study, is relatively small. Nevertheless, it still provides some value as it does confirm a trend picked up by other researchers: that of young adults devoting more time to online pursuits and less to reading.

For instance, in (Mokhtari, Reichard and Gardner, 2009), a majority of 85% of the students surveyed indicated internet use as their most favoured activity over television, academic reading and leisure reading. The demographic for this study was somewhat broader, with a random selection of all undergraduate students at the university recruited. Again, the majority age group amongst participants was early twenties, with a majority of 65.1% falling into the 20-22 age range.

Interestingly, Mokhtari et al. also found that in their test sample, there was no statistically significant relationship between the amount of time reported as spent on the internet and time spent reading or watching television, suggesting that whilst time spent online is growing, it is not necessarily the cause of the reported decline in time spent on reading. This lack of a provable causal link may primarily be due to the large amount of multitasking between the four measured activities (television, internet, academic/leisure reading) which the subjects reported. In this study, like in (Huang et al., 2014), time online was also treated as separate from reading activities, whilst in reality many of the students’ online pastimes, such as email and chatting, were text-based. This also contributes to the blurring of lines between reading and web usage, making “time spent reading” more difficult to measure.

This question was addressed more specifically by (Akarsu and Dar, 2014), who surveyed students on how much of their online time was spent reading, and what kind of materials they read. Their overall breakdown of online activity time is shown below:
Reading in general was fairly popular, with news being the most frequently read material, followed by email, weather reports and comics. Social media sites were not included as a category, but given the popularity of the “chatting with friends” activity, it may be assumed that as in other studies, these were also a popular activity.

Taken together, these studies do appear to confirm the perception that reading books for leisure, whilst still enjoyed by young adults, is in decline, with this age group’s primary text interaction most likely to be either social media or practical online information sources.

Finally, to consider a more general survey of adult reading preferences, the Booktrust Reading Habits Survey (Gleed, 2013) collected data from over 1500 adults in England. This survey was carried out using quota sampling to ensure that the distribution of participants was a match for the demographic profile of the country as a whole, and therefore provides a much broader picture of adults’ reading habits than those carried out with younger adults.

In the instance of young adults, however, the results vary little from the American college studies: whilst 27% of participants overall preferred going online to reading books, this rose to 56% in the 18-30 age group. With average literacy proficiency scores for young adults in the UK worse than those for the population as a whole (262 points on the OECD 500-point proficiency scale, as opposed to a national mean of 270) (OECD, 2013), this data would seem to confirm that both reading enjoyment and reading ability is indeed in decline amongst this age group. Overall, whilst younger adults’ reported attitudes towards reading are still fairly positive, there appears to be a split between theory and practice, with many saying that they do enjoy reading but spending less time actually doing so.
In terms of what types of books adults are reading, there is a more even split in preferences between fiction and non-fiction than in children and teens, with a 44% majority stating that they read a mixture of both. A breakdown of overall preferences by genre is shown below:

<table>
<thead>
<tr>
<th>What types of genres of books do you find most enjoyable? (Overall)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genre</td>
</tr>
<tr>
<td>-----------------------------</td>
</tr>
<tr>
<td>Crime/thrillers/mystery</td>
</tr>
<tr>
<td>(Auto)Biographies</td>
</tr>
<tr>
<td>History</td>
</tr>
<tr>
<td>Science fiction and fantasy</td>
</tr>
<tr>
<td>Romance</td>
</tr>
<tr>
<td>Historical fiction</td>
</tr>
<tr>
<td>Special interests / hobbies</td>
</tr>
<tr>
<td>Classic novels</td>
</tr>
<tr>
<td>Modern fiction</td>
</tr>
<tr>
<td>Horror</td>
</tr>
<tr>
<td>Humour</td>
</tr>
<tr>
<td>Sports</td>
</tr>
<tr>
<td>Science</td>
</tr>
</tbody>
</table>

Figure 3 - Adult reading preferences by genre (Gleed, 2013, Fig.25, p.9)

The Scholastic and BookTrust studies provide an excellent opportunity to compare data as far as genre preferences go, with both taken from population-representative samples (albeit one UK and one US), with a large sample size and the same format of question used. In both studies subjects were given a list of genres (the children’s options somewhat more generalised, but still largely comparable) and allowed to pick as many as applied. This gives comparable percentage data showing the change in preferences between children and adults, and shows that there is indeed a significant shift in preferences with age.

As Figure [3] shows, adults’ most preferred type of reading material is novels including an element of mystery or suspense, a taste which has jumped up several ranks from its middling position when rated by children (Scholastic Inc. and YouGov, 2015). Conversely, the runaway children’s favourite, comedy books, has dropped significantly in popularity, falling to the middle ranks amongst adults.

This difference in genre preferences will be highly significant and should be taken into account when designing materials aimed to motivate readers across different age ranges.

Having considered what adults read, there is also the question of what motivates them to read for pleasure, or indeed puts them off doing so. On the positive side, the primary motivations cited were escapism and relaxation (Gleed, 2013), much the same as (Dungworth et al., 2004) found with children.
For those who did not enjoy reading, their main reason given was that they didn’t have the time, whilst second place was a tie between having more interesting things to do and a sense of trying to read but quitting due to boredom/lack of patience. Again, this trend appears to be consistent throughout the different age groups – a sense of lacking the time for reading appears frequently in answers as to why people don’t read or read less than they once did. One theory which might be drawn from this is that shorter texts may therefore be more appealing to those who feel they lack the time to read a book – this may in fact correlate with the growing popularity of graphic novels reported amongst the college-student age group.

Given the proportion of the adult population who can be estimated to be only functionally literate, difficulty with reading was surprisingly low on the list of de-motivational factors – only 17% of respondents agreed that reading was “hard work,” placing it as the third most common reason, compared to 36% who “lost patience” and 35% who “didn’t have time”. This was the same when children were asked why they enjoyed reading less than when they were younger – harder books came fourth out of the five main reasons given.

Text difficulty might, however, be linked to the high incidence of boredom or inability to engage with the book that was reported – if the text is too complex to read without significant effort, this will certainly make it harder to engage with the content. This would fit with the proven link between reading ability and reading motivation which has already been discussed. If text difficulty is indeed a driving factor behind people being put off reading, providing leisure reading with simpler texts might in fact aid these reluctant readers in taking up reading.

To summarise, it appears that reading materials designed to motivate independent reading in children should be fiction-based, fun and humorous, whilst for adults, escapism and mystery are considered most enjoyable. For all age groups, since time available is commonly cited as a reason for not reading, content in short, episodic formats may be most successful, and level of language should not be inappropriately high so as to avoid presenting a barrier to engagement.
1.4 – How do eBooks relate to reading motivation?

With a rise across all age groups in the reported time spent using computers, smartphones or other electronic devices online, how do digital books fare against their paper counterparts?

Surprisingly, given the popularity of digital media, eBooks overall have been less than the runaway success once predicted to put paper bookstores out of business.

Beginning with usage of eBooks amongst children, figures show that more kids have read an eBook in recent years (Scholastic Inc. and YouGov, 2015), but those with a preference for digital books are still in the minority. 70% of the children surveyed in one international study expressed a preference for print books over electronic (Kleeman, 2016), and the Scholastic survey’s results concur, with 65% of subjects agreeing that “I’ll always want to read books printed on paper even though there are eBooks available”. Interest in trying eBooks amongst those who had not done so had also dropped recently, suggesting that the initial societal novelty factor of devices such as the Kindle first being released has worn off.

As for teenagers, a recent study carried out in Norway compared experiences with reading the same text in print and on a dedicated e-reader device with a group of 15-year-olds (Tveit and Mangen, 2014). Preference for the e-reader device was largely affected by subjects’ existing reading habits: those subjects who already identified as book lovers were least likely to prefer the electronic format, whilst less avid readers were more inclined towards the electronic format. It was particularly popular amongst certain subgroups, with the authors finding that: “reluctant boy readers valued reading more after their experience with e-readers.” (Tveit and Mangen, 2014, p.183)

These results concur with those of (Picton and Clark, 2015), who, when studying use of eBooks with 8-16-year-olds, also found that electronic books were more popular (47.6%, when measured at the project’s end, compared to 19.1% for paper) and had a positive effect on motivation. Over the course of the project, the number of pupils who rated their reading enjoyment as “very much” rose from 9.6 % to 18.3%, and the technology was mentioned as being exceptionally helpful to some reluctant readers.

Amongst young adults, studies found were primarily focused on academic reading rather than leisure reading, and found that in this situation, an interesting contradiction exists (Vandenhoek, 2013), (Mizrachi, 2015). Student preferences overwhelmingly favoured reading academic materials on paper, with the ease of the ability to underline and annotate directly on the text cited as the most common reason, yet most actually did their reading online, citing cost and environmental concern as their main reasons for not printing more.

In (Foasberg, 2014), which surveyed students on both academic and leisure reading, the trend of preferring paper books for academic purposes was again confirmed, but a strong
contrast between reading format preference by content type also emerged, as shown below:

![Figure 2 - Academic Use of Formats](image)

**Figure 4 – Academic Use of Formats, (Foasberg, 2014, Fig.2, p11)**

Here, it can be seen that digital reading formats were overwhelmingly preferred for personal reading, with mobile devices much more popular with this content type.

Amongst the general adult population, readers of eBooks are still in the minority, with overall format preference sitting at around 76% towards paper (Gleed, 2013). Surveys do, however, show a slowly increasing rise in ownership and use of both dedicated e-readers and tablets as reading devices (Rainie et al., 2012). Regular readers of eBooks were also found to be amongst the most frequent readers, consuming more books overall than traditional paper readers. Rainie et al. also found that 30% of their respondents who consumed digital books said that the availability of digital content meant they were reading more now than previously (62% read the same amount and 7% less). This was particularly true with younger male readers at 33%.

Overall, the uptake of digital books remains mixed, but the growth in ownership of tablets, smartphones and e-readers may predict an increasing rise in their usage. As pointed out in the above study:

“...people who specifically own e-reading devices (readers and tablets) are particularly likely to read e-books: 93% of e-reader owners consume e-books on their traditional Kindle or
Nook at least occasionally; 81% of tablet owners consume e-books on their tablet; 46% of computer owners consume e-books on their computer; and 29% of cell phone owners read e-books on their phone.” (Rainie et al., 2012, p.41).

EBooks (at least for leisure reading) appear particularly appealing to the teenage and young adult groups, especially to some more reluctant readers. This may make these groups the best candidates to target for promoting eBook-based reading materials.
1.5 – What About Interactive eBooks?

Interactive eBooks, a format defined by being more than just a digitised copy of a paper novel, are currently still a new, relatively little-used format. These books, usually either website-based or in the form of a mobile app, can include multimedia including video, images and sound, as well as user interactions of various types.

Some examples of highly-rated interactive eBooks are *Arcadia* by Ian Pears (iOS), which allows users to switch between interweaving perspectives and different characters’ storylines at will, or *80 Days* by Inkle studios, a modern update on the choose-your-own-adventure novel (Android & iOS) (Dredge, 2015).

Recently the most popular format for interactive eBooks (also referred to as eBook apps) has been Apple’s App Store (Sargeant, 2015), with children’s storybooks being one of the most popular categories.

As pointed out by Sargeant’s analysis of some of the more popular of these picture books, the interactive format is more engaging for children, but caution is needed in designing book apps so that interactivity enhances rather than distracts from the text.

This concurs with (Takacs, Swart and Bus, 2015), who conducted a meta-analysis on the use of enhanced eBooks in with young children and its effect on their literacy levels. Overall, their findings stated that whilst the inclusion of multimedia content was beneficial to subjects’ enjoyment and literacy levels, the use of interactive features such as minigames and dictionary functions was found to be a distraction from engaging with the story.

For older readers, research on the interactive eBook appears to be largely untapped, with little information available on their uptake or effect on reading habits. This may be due to developers’ focus on children’s books, with the enhanced picture book the most commonly available form of interactive eBook at this time.

The growth of interactive eBooks as a format has also arguably been hindered by the fact that there is no single unified format which can be read across all mobile devices, as pointed out in (Constanzo, 2014). Some progress towards this has been made, with the EPUB3 standard developed by the International Digital Publishing Forum (Conboy et al., 2011) published in 2011 and approved by the ISO in 2014.

This standard was developed to replace the EPUB format, an open file standard which is currently the most widely-used format for eBooks, supported by the majority of e-reader and mobile devices. It supports only text and image content, and therefore was ideal for traditional eBooks as a simple analogue of print content, but lacked support for anything more. EPUB3 includes provision for multimedia and interactive content, but whilst this new standard is available, at present online bookstores still use EPUB for their content, and presumably will continue to do so until more EPUB3-specific content is authored. Until such
a time, the platform-specific app (primarily iOS or Android) will likely remain the most popular source for interactive eBook content.
1.6 – Interactive Storytelling in Research

Whether in the form of eBooks, video games or mixed-media apps, the subject of interactive storytelling has been considered extensively in research, and as such, should be considered when designing an interactive storybook.

As pointed out in (Riedl and Bulitko, 2013), the central problem of interactive storytelling is: “how to balance the need for a coherent story progression with user agency” (Yu and Riedl, 2014, p.61). In other words, how can a system generate content dynamically, allowing user interaction, whilst still forming a story that makes sense?

In an attempt to address this problem, a spectrum exists between pre-authored content over which the user has no control and auto-generated content with a high degree of user input. A traditional text-only eBook would sit at the very end of the scale for the former, whilst an Experience-Manager based interactive story, in which the player influences progression at every point, would occupy the other end. One representation of this scale showing the relative positions of some existing systems, by (Riedl and Bulitko, 2013) is shown below:

![Figure 5 - (Riedl and Bulitko, 2013, Fig.2, p71)](image)

This section will briefly consider how existing systems address the issue of balancing interaction and control, and which of these systems might be most suited to interactive eBooks.
An overview of the different types of interactive story structure which have previously been developed is given in (Luo et al., 2015), in which the authors examine nine interactive narrative systems as a representational sample of the current state of the art.

In an auto-generated story, the user controls a virtual character and influences the development of the storyline through their actions. In systems with auto-generated storylines, some form of artificial intelligence approach, whether agent-based or state-based, is generally required in order to make decisions on how to generate the next piece of content.

In some systems, such as Thespian (Si, Marsella and Pynadath, 2005), virtual agents are used to represent characters. These interact with one another and the player agent in a goal-driven manner defined by the author, as well as by a director agent, which can influence the other agents to make sure the generated story stays within the bounds of the system.

The “boundary problem,” as discussed in (Riedl and Bulitko, 2013), is an issue which arises in auto-generated systems – how to keep the plot from straying outside the set of conditions which the system can cope with. Like Thespian’s director, most auto-generative systems therefore employ some form of Experience Manager, a software agent which guides progression by either preventing or accommodating out-of-bounds actions.

Other systems, such as Mimesis (Young and Riedl, 2004) and Automated Story Director (ASD) (Riedl et al., 2008), use a planning-based narrative generation approach. These are state-based systems, with a set of pre-defined actions available and a story world with a set of current and goal states. Content is then generated according to possible actions given current states.

In addition to purely auto-generated content, there also exist what Riedl and Bulitko call “hybrid systems,” in which there exists a greater degree of manual author control over how the storyline progresses. For example, the narrative management structure of Façade (Mateas and Stern, 2003) (which divides story elements into micro-units termed “beats”) is outlined below:
In this system, the author defines the desired story arc, which the Drama Manager (another term for the Experience Manager) then adheres to when selecting the next story element to progress to. In this way, content is both auto-generated and author-constrained.

The PaSSAGE system (Thue et al., 2007) takes a different approach to constraining story flow, and employs a tree-based structure, also known as a branching story graph, to determine narrative progress. The player’s influence is limited to set decision points encountered at specified points in the tree, as shown below:
Some more advanced versions of the branching graph model (PaSSAGE, PCBF system, (Yu and Riedl, 2014)) also employ an experience manager which collects data on the player’s preferences as they progress and presents appropriate next options based on this model in order to create a more personalised experience.

Each different approach to interactive content allows for a different level of controllability and interaction – at the core, how much the player’s actions influence the story. Each has its own benefits and pitfalls in terms of user engagement and enjoyment of the content.

The question of which approach to take when designing interactive eBook software then comes down to prioritising which aspects of the user experience are most important. As suggested in (Bruni and Baceviciute, 2013), doing this in the design of an interactive system requires defining the overall goals of the system.

Given the survey data previously discussed regarding readers’ motivations for choosing books, an enjoyable story was featured frequently as a central requirement. With the project’s overall aim of increasing users’ time spent reading, engagement is also a key goal. Taking enjoyability and engagement as the system’s key goals, it is then possible to evaluate approaches.

Users evaluating the Façade interactive system in (Seif El-Nasr, Milam and Maygoli, 2013) found that their experiences were detracted from by some NPC responses breaking suspension of disbelief by not responding appropriately to player input. Timing and pacing was also an issue for some users.
Narrative coherence may be the greatest problem when considering the use of auto-generated story elements; as discussed in (Green and Jenkins, 2014), it is difficult to create a meaningful narrative without a certain level of author control to allow affective storytelling. The problem of the inability of auto-generative systems to consistently produce “compelling” content is also highlighted in (Aylett, Louchart and Weallans, 2011). Enjoyment of the story may therefore be better served by having more author control over content.

Some form of hybrid model may therefore be more appropriate, in allowing the user some sense of agency in order to increase their engagement (Bruni and Baceviciute, 2013), whilst maintaining enough author control to ensure a smooth story flow.

One such model which achieved reasonable commercial and critical success was an attempt at the “interactive movie,” Quantic Dreams Studio’s game Heavy Rain. The game uses a branching narrative structure with a relatively narrow set of available options, allowing events to progress in a believable way, yet also permitting enough player influence for choices to feel meaningful. A representation of one chapter of the narrative structure, taken from (Wei, 2011), in which the game was analysed as an example of interactive narrative, is shown below:

![Figure 8 – Heavy Rain narrative structure, (Wei, 2011, Fig.2, p.3)](image)

As Wei highlights, Heavy Rain is able to engage players, giving them enough variety in terms of story pathways to make replaying viable and enough influence to feel a measure of control. This is despite the fact that its fairly restrictive structure also allows for more author control in order to maintain a coherent plot. The balance achieved in this model is close to the ideal state desired for the interactive eBook app, and suggests a possible outline on which to base the narrative structure.

Given both the overall aims of this project and the practical concerns of authorship, from this analysis it would appear that a branching-story-tree-based model is the most practical
form of story structure. This can allow user interaction in the form of path choice, much like the choose-your-own-adventure novels used as models in several of the research projects discussed, whilst maintaining a reasonable scope and allowing a superior degree of manual author control.
1.7 – Summary - The Story from Here...

From the research considered, it is clear that the overall decline in literacy levels in the UK and other developed nations is related to the trend of reading for entertainment less often. In particular, time spent reading declines markedly in children around the age of 8, and again in young adults (age 18-30).

Rather than competing with the digital media which now takes up more of our leisure time, adapting to the electronic platform may therefore be the best strategy for promoting reading. To take advantage of the full capabilities of mobile devices, in terms of providing multimedia content, developing an interactive eBook using the mobile app format appears to be the best way forward.

To promote users feeling engaged with the story, a degree of user influence will be included, based on the branching-story model discussed earlier. This allows a feeling of user agency whilst allowing the majority of the content to be authored in a controlled manual fashion.

The development and evaluation of such an app will face several challenges: for instance, the evaluation of concepts such as enjoyment and engagement is notoriously subjective and difficult to quantify. This therefore requires careful experimental design, which will be addressed further in the next section.
2 – Requirements Analysis

2.1 – Project Brief & Mission Statement

The initial brief for this project (reformatted for clarity) reads as follows:

“eBooks are mostly very simple and do not support:

- Differing levels of linguistic complexity
- Differing richness of non-textual materials (pictures, cut sequences etc.)
- Differing levels of detail overall, for particular characters etc.

Nor do they support:

- Random or controlled interleaving of pathways
- Selection of alternative endings.

This project will create an app which supports all these features and people it with a selection of fairy stories.”

From that problem statement, a mission statement was developed:

This project will create a fully functional eBook app with the following features:

- A mixture of multimedia content
- A choice of pathways through the story
- Multiple levels of detail/difficulty

The aim of creating this app is to test whether an interactive storybook app can provide increased user engagement and enjoyment, especially for poor readers, compared to a traditional paper novel or text-only eBook.
2.2 – Stakeholder Requirements

To begin, an assessment of who the project stakeholders would be was conducted in order to determine what their requirements were likely to be.

The demographic groups chosen for testing were selected as those most likely to benefit from using the app and are as follows:

- Children aged 7-9 – this is the average age at which children start to lose interest in reading
- Young Adults (18-30) – this demographic group typically reads less than any other

Therefore, for the purposes of this project, direct users would fall into these two groups. Secondary users would be parents, schools and teachers in terms of the age 7-9 user group. Other stakeholders are the developer and the University.

Considering these user groups, the following list of requirements was created:

Everyone:

- Simple, easy to use interface
- Works across Android devices of a variety of specifications
- Entertaining humorous content
- Accessibility support features (considering dyslexia, colour blindness)

Kids 7-9:

- Short, easy-to-digest content chapters
- Age-appropriate content and difficulty
- Bright, engaging design style

Parents, schools and teachers:

- App does not require internet access – safe and secure for unsupervised use
- Easily see user progress from the map menu
- Maintain text to multimedia ratio in favour of text to ensure users do an adequate amount of reading

Young Adults 18-30

- Short content chapters that can be fit into a busy schedule
- Different depth/complexity of text depending on difficulty selection

Developer/University

- App looks and behaves like a professional piece of software
- App demonstrates a variety of features and design/programming skills
2.3 - Software Requirements

Based on the mission statement, software requirements were divided into three main categories: non-linear pathways, multimedia materials, and multiple levels of detail/difficulty.

Phase 1 of the software requirements analysis was carried out by breaking down the project brief into a concept map, expanding from general statements into specific goals, as shown below in Fig. [9]:

From the concept map, a detailed list of project objectives was developed, with additional
technical objectives not covered in the initial map added as necessary.

Objectives were grouped into sections according to which overall aims they fall under, and each was assigned a priority level. Items assigned red priority are considered critical, and these goals must be met for the project to be considered complete. These largely relate to functional program objectives required for the app to run and be usable for user testing.

Those assigned amber priority are expected features, but are not considered absolutely essential. Those marked in green are stretch goals; features over and above the basic project specification which will be added as time and resources allow in order to provide additional functionality.

This list became the project checklist V1.0, as shown in Fig [10] below.

<table>
<thead>
<tr>
<th>Selection of multimedia content</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Text</strong></td>
</tr>
<tr>
<td>• Displays text in flippable pages like a standard eBook reader</td>
</tr>
<tr>
<td>• Fetch text from file and split to pages dynamically depending on page/font size</td>
</tr>
<tr>
<td>• Remembers reader progress on resume</td>
</tr>
<tr>
<td>• Allow user to choose text properties (size, colours, font set)</td>
</tr>
<tr>
<td>• Column view for larger landscape devices</td>
</tr>
<tr>
<td><strong>Video</strong></td>
</tr>
<tr>
<td>• Play short video clips</td>
</tr>
<tr>
<td>• Include media controls for play/pause etc.</td>
</tr>
<tr>
<td>• Buffering % and load bar display for slower devices</td>
</tr>
<tr>
<td>• Optimise layout for video display on different devices</td>
</tr>
<tr>
<td>• Produce 3d animated content for clips</td>
</tr>
<tr>
<td><strong>Images</strong></td>
</tr>
<tr>
<td>• Flippable image viewer</td>
</tr>
<tr>
<td>• Option for display as slideshow or manual page turn</td>
</tr>
<tr>
<td>Audio</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>• Background audio loop for images/text</td>
</tr>
<tr>
<td>• User can select track, enable/disable sounds</td>
</tr>
<tr>
<td>• Touch-to-hear speaking dictionary function on text select</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Interactive elements (minigame)</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Small minigame “quick-time event” elements to influence available story paths/outcomes</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Voice-over/ storyteller</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Option to have text with synced voice-over and highlight</td>
</tr>
<tr>
<td>• Storyteller mode with auto-play, user can set up and have story played for them</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Non-linear pathway options</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Branching story nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Map-based node menu presents available content and structure visually, showing difficulty level and description for each node</td>
</tr>
<tr>
<td>• Each node can contain a playlist of mixed-media content which will switch media players seamlessly as the user progresses through the set</td>
</tr>
<tr>
<td>• Nodes unlock depending on user mode setting and progress</td>
</tr>
<tr>
<td>• Unlock progress saved between sessions</td>
</tr>
<tr>
<td>• Multiple paths available with different outcomes</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Different unlock modes</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Choose-your-path mode, unlocking as the player completes each node, with further nodes opened depending on previous choices</td>
</tr>
<tr>
<td>• Traditional mode, with all content presented in a preset order</td>
</tr>
<tr>
<td>• Freestyle mode, with all content unlocked from the start, user can pick and choose with no need to follow a chronological path</td>
</tr>
<tr>
<td>• Hidden nodes, available only when a certain node set has been unlocked</td>
</tr>
<tr>
<td>• Random bonus content nodes, which may insert themselves at the start/end of the selected node under preset conditions</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable levels of difficulty</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Nodes with different difficulty ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Provide easy/medium/hard text levels</td>
</tr>
<tr>
<td>• Rate text using standard readability testing scales</td>
</tr>
<tr>
<td>• Display difficulty rating on node map menu</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Core and non-core story paths offering choice of length/depth</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Non-core paths with higher difficulty ratings available to choose without being necessary to progress</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Different versions of core path text depending on selected difficulty level</th>
</tr>
</thead>
<tbody>
<tr>
<td>• User can select overall difficulty level, which will display core text versions capped at the selected difficulty</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dynamically alter core text depending on difficulty</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Alter core text complexity algorithmically, rather than manually including 3 versions in the source text files</td>
</tr>
</tbody>
</table>
App structure

App must be simple and easy to use
- Interface must be familiar and intuitive to use
- Android Material Design specifications used wherever possible
- User settings profile auto-saves user progress and settings between sessions
- One simple settings dialogue available from anywhere in the app to change settings

App should be modular to support expansion
- Modular node structure to allow for easy insertion/removal of node content
- Overall modular Book class to allow loading and adding of multiple books consisting of multiple nodes using the same app code.

Figure 10 – Project Aims, Version 1.0

This initial project checklist was then used as a guideline for planning the software development process, along with feedback from the initial user consultation.
3 – Methodology

Derived from the mission statement shown previously, the primary research question for this project is:

“Can an interactive storybook app provide increased user engagement and enjoyment, especially for poor readers, compared to a traditional paper novel or text-only eBook?”

This can be broken down into the following hypotheses:

- An interactive storybook app is more enjoyable than a traditional book/eBook
- An interactive storybook app is more engaging than a traditional book/eBook
- An interactive storybook app is more motivating in terms of reading more often than a traditional book/eBook
- Users who find reading difficult/uninteresting would read more often using an interactive storybook app

In order to test these hypotheses, an appropriate app would have be developed and tested with the chosen user groups to assess its effectiveness.

Following the research on interactive storytelling software previously discussed, the Android app format was chosen as the software development platform for this project. This was for several reasons:

- Developer familiarity with the basics of the primary coding languages (Java/XML)
- Easy access to devices for testing
- Rich inbuilt libraries for multimedia handling
- Unlike iOS, Android apps are reflow-able (able to be used on devices of different screen sizes/formats without releasing separate app versions), allowing wider usage
- Android currently has the largest smartphone market share
- Media files can be bundled directly within the app, unlike a web-based solution, meaning the potential problems of accessing internet resources during testing are negated

The EPUB3 eBook format was also considered, but was rejected in favour of Android app due to current lack of widespread format support and the fact that the file structure is designed for linear rather than branching content.

The project was modelled on an iterative user-centric design model, with user consultations taking place at three stages of the development process. This would help ensure that the resulting app would be something users could and would want to actually use.

A mixed-methods approach was selected for analysis, collecting a range of both qualitative and quantitative data. The formats chosen were:
• **Initial consultation** – informal online questionnaire featuring an app design mockup video (an interactive PDF file version of this mockup can be found in the Appendices). (Likert scale and open-answer boxes)

• **Experiment 1** – User app testing followed by online questionnaire (Likert scale and open-answer boxes)

• **Experiment 2** – User app testing followed by focus group discussion (semi-guided structure)

This combined methodology was chosen as it is difficult to present a clear picture of such a subjective topic based on one data type alone. The combination of hard data from rating scales and detail from open-comment responses and focus group discussion should give a more comprehensive result.

These methods were also chosen due to being appropriate to the timescale of the project. This time limitation influenced the choice of tactics in two major ways: the decision to split user evaluation into two experiments, and the choice of using self-reported responses to measure enjoyment and engagement.

The former decision was mainly influenced by the timeframe available for conducting user testing in schools – as the most viable method of accessing groups of children for testing with younger users, school testing is an important stage in the evaluation process. Due to the structure of the school year, this meant conducting testing before the end of June, relatively early in the project’s schedule. Therefore, Experiment 1 would run with an initial, basicFeatured prototype in order to obtain the first batch of user feedback in time. Experiment 2, using young adult participants, would be conducted with a more advanced version of the prototype later in the project to allow testing of a more full-featured version.

The second concern brought about by timeframe is the choice of methodology for measuring users’ responses in terms of engagement and motivation. These qualities are difficult to measure, a problem encountered in research regarding users’ feelings and behaviours.

The majority of the studies referenced earlier in this paper addressed this problem through two main strategies: long-term behavioural studies and self-reported scores.

As the timeframe for this project is not suitable for long-term behavioural-data-based studies, evaluation will therefore have to rely largely on self-reported measures. This will consist of asking subjects whether they think they feel more engaged/would read more often based on only one session using the app. Whilst this is not ideal, as it is based on users’ own estimates of how much they are affected rather than actual usage data, it should still produce enough data for comparative purposes and may pave the way for further study.
4 – Project Plan

Following the definition of project aims and selection of methodology, a structured project plan laying out goals and timeframe was set out at the beginning of the project to serve as guidance for progress.

The project was divided into six major stages, each consisting of several sub-goal blocks in an Agile-type arrangement. A Gantt chart (below) was used to define an estimated timeframe for each stage. This schedule and structure were followed throughout the project, with only minor timetabling and methodology changes made along the way.

<table>
<thead>
<tr>
<th>Task</th>
<th>Start Date</th>
<th>Duration</th>
<th>Q2 Jan</th>
<th>Q2 Feb</th>
<th>Q2 Mar</th>
<th>Q2 Apr</th>
<th>Q2 May</th>
<th>Q2 Jun</th>
<th>Q2 Jul</th>
<th>Q2 Aug</th>
<th>Q2 Sep</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 1</td>
<td>11/01/16</td>
<td>81d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic App</td>
<td>20/01/16</td>
<td>72d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Video &amp; Image viewer modules</td>
<td>20/01/16</td>
<td>6d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test Viewer module – Basic</td>
<td>27/01/16</td>
<td>69d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic menu &amp; function integration</td>
<td>21/02/16</td>
<td>10d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic Function complete</td>
<td>01/03/16</td>
<td>64d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial Consultation</td>
<td>16/03/16</td>
<td>16d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Background Research</td>
<td>11/03/16</td>
<td>72d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Background Reading &amp; Source Collection</td>
<td>11/03/16</td>
<td>47d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Research Report Draft</td>
<td>01/02/16</td>
<td>20d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Research Report draft complete</td>
<td>26/02/16</td>
<td>0d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Research Report final copy editing</td>
<td>27/02/16</td>
<td>54d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Research Report hand-in</td>
<td>30/03/16</td>
<td>0d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage 1 Complete</td>
<td>31/03/16</td>
<td>0d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage 2</td>
<td>06/04/16</td>
<td>94d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Node class development</td>
<td>06/04/16</td>
<td>94d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>User Profile App data development</td>
<td>13/04/16</td>
<td>15d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>App Structure Complete</td>
<td>27/04/16</td>
<td>0d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test content creation &amp; tailoring</td>
<td>28/04/16</td>
<td>94d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multimedia content creation</td>
<td>04/05/16</td>
<td>15d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prototype 1 Testing</td>
<td>20/05/16</td>
<td>15d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prototype 1 complete</td>
<td>06/06/16</td>
<td>0d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage 2 complete</td>
<td>06/06/16</td>
<td>0d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage 3</td>
<td>08/06/16</td>
<td>14d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment 1</td>
<td>06/08/16</td>
<td>94d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment 1 Analysis</td>
<td>10/08/16</td>
<td>74d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage 3 Complete</td>
<td>21/08/16</td>
<td>0d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage 4</td>
<td>23/08/16</td>
<td>22d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prototype 2 Development</td>
<td>23/08/16</td>
<td>22d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage 4 Complete</td>
<td>13/07/16</td>
<td>0d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage 5</td>
<td>13/07/16</td>
<td>15d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment 2</td>
<td>13/07/16</td>
<td>94d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment 2 Analysis</td>
<td>20/07/16</td>
<td>94d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage 5 Complete</td>
<td>27/07/16</td>
<td>0d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage 6</td>
<td>27/07/16</td>
<td>30d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final Analysis &amp; Report Writing</td>
<td>27/07/16</td>
<td>15d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Draft Final Report</td>
<td>10/08/16</td>
<td>94d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final Project Submission</td>
<td>10/08/16</td>
<td>0d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poster Development</td>
<td>10/08/16</td>
<td>74d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Project Poster Presentation</td>
<td>22/08/16</td>
<td>1d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Project Complete</td>
<td>25/08/16</td>
<td>0d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 11 – Project initial Gantt chart
A summary of each stage and its aims and progress, based on the original plan, is given below:

**Stage 1 – Initial Consultation & Development**

This stage consisted of three activities:

- **Background research (presented earlier in this paper):** research and analysis of the current state of the art in the areas of literacy, reading motivation and interactive storytelling.
- **Initial software design/development:**
  - Development of basic Android app framework to build familiarity with the development environment
  - Rough planning of class structures
- **Initial user consultation:** a small-scale open user survey was conducted to ask users:
  - Which of the proposed optional software features (outlined previously in Project Aims) they would rank as most important/attractive.
  - What they thought of a basic mockup of the proposed interface design in the form of a short demonstration video.

Stage 1 was completed on time without any major difficulties.

**Stage 2 – Prototype 1 Development**

This phase focused on the development of the basic app prototype, culminating in the Alpha build which included the following features:

- Text viewing (page-style reflow text based on device screen size)
- Image viewing
- Video viewing
- Branching content structure accessed through a central menu
- Three textual difficulty modes
- Settings menu with options for sound, text size, difficulty mode

Stage 2 was finished on time, but it was necessary to drop one planned feature from the app: background sounds for text and image pages. This feature was removed due to the large time input required to create sufficient sound content versus the relatively small added value it would bring to the app.
Stage 3 – Prototype 1 Testing and Analysis

Prototype 1 was tested in classroom sessions in a local primary school. A total of 107 participants took part, well over the initial aim of 60 subjects.

The experiment used the following format:

Children were given around 30 (raised from the initial estimate of 15) minutes to read and interact with the storybook app.

At the end of the reading session, subjects were asked to fill out a short survey. This survey included the following:

- **Demographics**: Participant age and gender, reading level and dyslexia diagnosis status (yes/no) if known
- **Quantitative feedback**: Five-point Likert scale questions evaluating enjoyment, engagement, whether subject would be likely to read more using such apps
- **Qualitative feedback**: Open-text question box asking for any comments on what users enjoyed, didn’t enjoy and would change or add.

Due to data sensitivity, surveys were completely anonymised.

Despite technical difficulties which will be discussed further in Section 6.1, and resulted in the need to use paper rather than online surveys, the experiment was completed successfully on the assigned date. There was, however, a delay of approximately 2 weeks in receiving the resulting data, which impacted on Stage 4 and beyond.

Stage 4 – Refinement and Expansion into Prototype 2

Stage 4 consisted of refining and expanding the prototype app created in Stage 2 for further testing in Experiment 2. The initial aim for this stage was to add some of the additional software features, such as minigame elements and audio, but given user feedback and the time available, focus was shifted to improving existing features. Key updates included:

- Stability fixes
- Automatic progress saving
- Expanded content
- Content unlocking structure and difficulty level system refined
- Map UI redesign with improved aesthetic and clearer navigation
- Concealed “Easter Egg” chapter included to demonstrate hidden node capability

Stage 4 was completed approximately 2 weeks after original target time.
**Stage 5 – Prototype 2 Testing and Analysis**

This stage of the project was hit by repeated delays and difficulties centring on finding and scheduling volunteers to take part in the evaluation. After struggling to match volunteers’ schedules to enable a group meeting within the time available, it was decided to allow them to complete the evaluation individually and post their answers via a series of open questions online.

Stage 5 was eventually completed with 5 responses, approximately 10 days after the target time.

**Stage 6 – Final Analysis and Report**

The final stage of the project consists of the writing of this report, which evaluates data from both experiments together to determine any overall trends in user opinion, whether the research questions have been fully answered, limitations of the project and any recommendations for further research.

Stage 6 was successfully completed on time.
5 – App Development

The app development process took place concurrently with other project activities and spanned almost the entire January-August period.

Development was carried out using the native Android Studio development package on a Windows PC, with a selection of Android devices (Samsung Galaxy S5 Mini, Samsung Galaxy Tab S, Archos 80 G9) used for on-device testing throughout. Coding was done primarily in Java, plus XML for UI layout. The app was programmed using a minimum Android API requirement of 11 (Android 3.0 “Honeycomb,” released 2011) in order to support as many older devices as possible whilst still having access to key features of the modern Android architecture.

The app was designed based on a modular structure (a basic outline design is shown below) using Android’s Activities framework. This modular system allows for content nodes to be easily modified, added or removed. It also allows the code to be used to create new interactive books, and leaves the way open for expansion to support multiple titles.

![Figure 12 - Early App Structure Model](image)

The central content structure was also chosen early in the design stage; it was decided that the idea of branching content choices could best be represented using the idea of a map, with a literal choice of paths controlling how the story unfolds. At this stage the story to use was also picked — as the brief specified fairy tales, the story of *Little Red Riding Hood* was chosen. This story was selected because it was based around a journey, which would allow it to fit nicely with the map concept.

A concept structure sketch from the early design phase is shown below:
Before beginning implementation, a virtual mockup (created using Balsamiq mockup software and available in interactive PDF format in the Appendices) was created to get an idea of the overall app structure and user interface (UI) style.

This was shown to potential users in the initial consultation via a demonstration video. Whilst there were only a few respondents, comments were largely positive, with the app design rated as clear to follow and easy to use. After receiving this largely positive response, this design was carried forward as the basis of the app UI.

Once an outline of the app and content structure had been constructed, software development was begun. The development process was broken down into individual goals which were then assigned a block of time, typically one or two weeks, in an Agile-style flexible management system. An example of a development block was that creating the user preferences and associated dialogues was given two weeks’ time allocation.

This allowed for regular self-checkpoints to gauge the project’s progress was on track but was not too restrictive when blocks needed to be rearranged or extended.

The UI style chosen was largely centred on Google’s Android Material Design specifications, a design language intended to create unity of appearance and UI function across all Android apps (Google, 2015). Material standard interface components were used throughout, and custom elements such as the icons in the app’s map screen were designed to match this style.
Software development was firstly focused on the mechanics of the app: creating the three media players (book-view text, video and image slideshow) and linking them together in a branching playlist system. Placeholder content was used at this stage to ensure the software worked without spending time on content development.

Once this system had its basic functions in place, development moved to user customisation, this being a key feature of the app as much as the branched structure. A settings dialogue allowing the user to modify the appearance of the app (change theme) and the text (colour, size, font) was added. Font options included Dyslexie, a font specially developed by Christian Boer (Dyslexiefont.com, 2015) for ease of reading for dyslexic users. User settings were automatically saved to a persistent preferences file to be remembered across app instances. Saved progress was purposely not included in the user settings file but rather left to be dealt with separately, so that resetting the UI would not also cause a user to accidentally lose their progress.

The last development stage for the initial app was content creation. This involved three types of media: text, images and video.

The text was based on the Grimm Brothers’ “Little Red Cap,” taken from D. L. Ashliman’s compiled translations (Ashliman, 2013). This was then modified and expanded by the researcher with original text. Text source files were graded for difficulty based on a combined score from the consensus readability scoring tool at readabilityformulas.com. This takes eight widely used readability formulas including the Flesch-Kincaid Grade Level (the most commonly used standard) and produces an aggregated score as well as showing each individual result.

To ensure an appropriate range of difficulty levels for Phase 1 testing, the difficulty levels were based on the average reading level of an eight-year-old, with Grade 3 (age 8-9) therefore set as the medium difficulty. Grade 1 (age 6-7) was used for the easy setting and Grade 5 (10-11) for hard mode. Text was mostly written at the hard mode level and simplified accordingly for the lower levels.

Images were mostly created using stock images from online repositories (shutterstock.com, 123rf.com, freepik.com). These were then modified and compiled into the app’s illustrations by the researcher. Illustrations in a flat, geometric style using bright colour were chosen as they fit with the Material Design UI style and would be engaging and child-friendly. Some still frames drawn from videos used were also included.

Videos were sourced online from YouTube.com, with one exception which was filmed by the researcher. A list of links for video and image sources is provided in the Appendices. Issues relating to fair use and copyright law regarding this are discussed more fully in the document entitled PLES (Professional, Legal, Ethical and Social issues), which consists of a full review of these concerns and is also included in the Appendices.
Content structure was changed somewhat from the initial idea of core and non-core paths, and two different structures were tested:

In the first Alpha build, all chapters (content nodes) were accessible at all difficulty levels, and had appropriately-levelled texts for each.

In the Beta build used in the second experiment, certain nodes were only visible/unlockable at higher levels, to encourage users to try out the harder level whilst giving easy-level readers a smaller, more manageable content selection.

Following Experiment 1, the app was also developed further to fix some stability issues, improve UI appearance and usability and add additional content.
6 – Experiment 1

6.1 – Implementation

An Alpha prototype build of the app, given a working title of Wander, was used for Experiment 1. It featured one variable-branching storyline with three text difficulty levels, video and image content and user-editable font and colour settings.

The experiment was carried out on 16th June with 4 classes of pupils from Echline Primary School.

Due to the school’s use of composite classes, the classes visited consisted of P3/4, P4, P5 and P5/6. Rather than restrict participation to only those class members at the Primary 4 and 5 levels in the two composite classes, the whole class was invited to participate. This resulted in a slightly wider range than initially estimated, with a total of 107 participants with ages ranging from 8 to 11. This number far exceeded the initial goal of 60 test subjects, allowing for a richer depth of information to be gathered.

A 45-minute session was conducted with each class, consisting of a brief introduction to the project, followed by distribution of tablets around the class. A total of 19 Android tablets consisting of a mix of Lenovo and Nexus devices were used. Pupils therefore either worked in pairs or individually according to class size. Sharing and passing-around of devices was encouraged to ensure each participant had a fair share of time using the app.

Due to difficulties encountered with device authentication on the school’s WiFi network, the original plan to conduct the follow-up survey via app was set aside. The backup plan of using paper survey sheets was implemented in order to avoid postponing the testing session. (A copy of the survey sheet is included in the Appendices). These were distributed to class teachers the following day and collected when completed.

Accordingly, the testing sessions consisted of approximately 30 minutes per class spent interacting with the app. Participants were given their tablets with the app’s main menu already running and allowed to explore freely, with no further instruction other than that included in the app’s introduction chapter. The researcher was available throughout sessions to answer questions and deal with any issues.

Participants were then issued their feedback sheets in a follow-up session in class, and the completed surveys were then collected for analysis.
6.2 – Results – Likert Data

Despite technical problems which initially disrupted testing, the sessions progressed well. Informal feedback given on the day was largely positive, with several participants asking when/whether the app would be available for their own devices.

Once collected, data from the survey sheets was input into digital format; an Excel spreadsheet for demographic data and Likert scale scores, and a Word document for any comments made in open-answer boxes. Of all of the responses, only one had to be discarded as incomplete, giving 106 valid response sets.

Analysis of this data is broken down below:

- **Demographics**

  Distribution of ages was somewhat skewed, with the majority of participants aged 8 (40) or 9 (41) (76.4% total). 23 were aged 10, and only 2 participants were aged 11, too small a group for any meaningful testing; thus, ages 10 and 11 were combined into a single group of 25 participants (23.6%) for all further analysis.

  The gender split amongst subjects was almost even, with 51 female, 54 male and one who chose the option of “Other/rather not say.” (The neutral option was made available both to support gender inclusivity and to provide a privacy option for those who did not want their gender recorded). With less than 1% difference in a group of over 100 (disregarding the single neutral response), this could therefore be reasonably called a balanced group in terms of gender.
- **Reading Capability**

Due to the fact that testing was conducted at the end of the academic year, teachers did not have up-to-date reading age scores or similar metrics available regarding their pupils. Accordingly, since there was not time to administer a reading-level test during the session, subjects were asked to report on their own reading capabilities to try and gauge an estimate of their skill levels.

This produced an interesting distribution of results, (shown left in Figure [15]) with all but 2 (98.2%) rating their reading level at *Fair* (#3 on the 5-point scale) or above. A majority of 55 respondents (51.9%) rated themselves as *Good* (#4), with 36 self-reporting as *Very Good* (#5, 34%). The categories of *Poor* and *Very Poor* got one response each.

![Reading Confidence](image)

Figure 15 – Experiment 1 reading confidence scores

Whilst these results show a very positive level of reading confidence amongst the children surveyed, they likely say more about subjects’ *perceptions* of their skill levels than their actual capabilities. It seems highly unlikely that in a sample of 106 schoolchildren, only two would struggle with reading, given the overall percentage of adults in the UK who are only functionally literate (OECD, 2012). It may also be theorised that the skew in results for this question is due to a fear of looking bad or not wanting to admit to struggling with reading.

As such, these results provide some interesting insights, but are not likely an accurate representation of the distribution of reading skill levels amongst participants. This question will therefore be treated as additional information rather than statistically reliable data.

Participants were also asked whether they had a diagnosis of dyslexia. Only two subjects answered Yes, though there were also 6 *Prefer not to say* responses, suggesting the real number may have been higher. It was interesting to note that only one of the two subjects who gave their reading level as less than *Fair* was dyslexic.

- **App Ratings**

The first main section of this experiment was the five-question Likert ratings section, with questions corresponding to each of the research questions for this project.

This measured each of the following on a five-point scale:
- How much they enjoyed using the app
- How engaging they felt the app was
- Whether they would use an app like this in future
- If they would want to read more often if they had access to books on an app like this
- Whether they enjoyed using the app more or less than a paper book

Response options were given as standardised symmetrical Likert-type phrases. For questions 1-4, the scale was labelled as below:

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Very Little</td>
<td>A Little</td>
<td>Some</td>
<td>Quite a Lot</td>
<td>Very Much</td>
</tr>
</tbody>
</table>

For question 5 (was the app better than paper books) the phrasing was modified to better match the question, with available responses of:

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Much Less</td>
<td>Less</td>
<td>Same</td>
<td>More</td>
<td>Much More</td>
</tr>
</tbody>
</table>

The five-point scale was chosen over using a 7- or 9-point to make filling out the survey quicker and simpler, and to prevent extreme-score avoidance bias as much as possible. The use of an odd-numbered scale with a neutral option was considered more appropriate as, when measuring variables such as enjoyment, it is entirely feasible for subjects to have neutral feelings and the scale should reflect this.

Responses covered a wide spectrum of opinion, from some who completely disliked the whole experience to others who were highly positive. Response charts and analyses for each of the five questions are shown below:

**Enjoyment**

The enjoyment rating showed a fairly wide range of distribution, skewed somewhat in favour of the more positive options, with response #4 ranking as the mode at 28 replies. Response options #2 and #5 tied as second most popular with 25 subjects choosing each. Overall, there were more positive replies than negative, with a combined total of 53 (50%) respondents choosing #4 or #5, whilst 32 (30%) chose #2 or #1.
Whilst this is a mostly positive result, the relatively high percentage of subjects indicating that they didn’t really enjoy using the app certainly shows room for improvement in future iterations of the software. The open-comment feedback which will be discussed later in this chapter will provide key information on what exactly users disliked about the app in order to make further improvements.

**Engagement –**

Engagement ratings show a strong central tendency, with response #3 most popular at 27 replies, followed by #2 and #4 tied at 24 each. 70% of users’ responses fell into these three categories.

This rather average response shows that the app had only moderate success in terms of engaging users. This may have been influenced by the rather short and busy nature of the testing sessions. It may also suggest a central tendency bias could have been present on the subject of engagement.

**Future Use –**

When asked whether they would use an app like this in future, users’ replies fell into a far more binary distribution, with answer #5 as the mode and #1 as second most frequent. As Figure [18] shows, this yes/no divide was quite sharp, demonstrating that many subjects had formed a definite opinion on whether they would or would not use this sort of software.

Taking #1 and #2 as a combined “No” category, and #4 and #5 as “Yes” results in comparative totals of 36% for No and 48% for Yes. With a majority of subjects falling into the Yes category, this is a largely positive result; however, the sharp split in responses suggests that to improve on this figure may be challenging. The high number of users who chose #1 would likely require significant changes/improvements in order to change their
minds. Again, the consideration of open-answer replies, particularly the negatives, will be key to understanding the reasons behind this result.

**Reading More Often –**

For the question, “Would you want to read more often if you had access to books on an app like this?” option #5 was the clear leader, with 34 subjects (32%) selecting it. Amongst the remaining 68%, responses were quite evenly distributed around the centre of the scale, with #3 being the second most popular response at 21%.

This suggests that the majority of subjects found the app would have only a small effect on their reading habits, but a significant minority thought they would be far more likely to read using the app. This relates to results from previous studies (Picton and Clark, 2015), in which researchers found that there was a certain subset of subjects for whom the app format was far more beneficial to reading motivation, whilst it had a lesser effect on others. As previous research indicated that the makeup of this subset is typically influenced by reading level, it is difficult to identify whether members of the same demographic were the app enthusiasts in this study.

Combining categories as previously to compare positive vs negative responses yields 51% positive vs. 27% negative, which is promising, but this must be considered alongside the result distribution to give a clear picture of just how many subjects thought the app would increase their reading frequency.

**Better Than Paper Books –**

The results for question 5 show a very similar distribution pattern to those for question 4: the mode is again #5, with 32% of responses, whilst #3 is second most frequent with 22%. Again, this suggests that aside from the group who found the app much better, the majority of subjects
found that there was little or no difference in their preference of one format over the other.

Converting responses to a Better/Worse binary comparison, the results show 48% thought the app was better, whilst 29% considered it worse. This also fits with previous research (Kleeman, 2016), in which a significant percentage of children surveyed still expressed a preference for paper books over electronic, but with the difference that in this case, that group was still in the minority. This change might be indicative of the growing use of, and normalisation of reading on, mobile devices over the years between those studies and the present.

A summary of the descriptive statistics for each of the five questions is given in Figure [21] below:

<table>
<thead>
<tr>
<th></th>
<th>Enjoyment</th>
<th>Engagement</th>
<th>Future</th>
<th>Frequently</th>
<th>Vs_Paper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>3.37</td>
<td>3.02</td>
<td>3.21</td>
<td>3.48</td>
<td>3.37</td>
</tr>
<tr>
<td>Median</td>
<td>3.50</td>
<td>3.00</td>
<td>3.00</td>
<td>4.00</td>
<td>3.00</td>
</tr>
<tr>
<td>Mode</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>1.260</td>
<td>1.280</td>
<td>1.485</td>
<td>1.325</td>
<td>1.430</td>
</tr>
<tr>
<td>Variance</td>
<td>1.587</td>
<td>1.638</td>
<td>2.204</td>
<td>1.757</td>
<td>2.044</td>
</tr>
<tr>
<td>Skewness</td>
<td>-.205</td>
<td>-.008</td>
<td>-.205</td>
<td>-.296</td>
<td>-.296</td>
</tr>
<tr>
<td>Std. Error of Skewness</td>
<td>.235</td>
<td>.235</td>
<td>.235</td>
<td>.235</td>
<td>.235</td>
</tr>
</tbody>
</table>

Figure 21 – Experiment 1 Results Descriptive Stats SPSS

Confidence intervals of the results for each of the five questions were also calculated, and can be found in the Appendices.

In addition to each individual question, a cumulative score consisting of the sum of each subject’s five ratings was also calculated. This provides an overview of the subjects’ general level of positivity towards the app.
As illustrated above in Figure [22], in general there was a moderate trend in favour of positive responses, with a median total score of 20/25 available rating points. The wide distribution of scores confirms the overall impression of mixed opinions with a positive skew.

### Data Relationships

Following the initial investigation of results, the data was cross-checked to determine whether any significant relationships exist between different user groups and their results. As Likert question data is non-parametric, statistical testing methodology was chosen accordingly. The Pearson Chi Square test was used to test for significance, and the resulting $p$ values are summarised in Figure [23] below:

<table>
<thead>
<tr>
<th></th>
<th>Enjoyment</th>
<th>Engagement</th>
<th>Future</th>
<th>Read More</th>
<th>Better</th>
<th>Combined Score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td>0.868</td>
<td>0.382</td>
<td>0.138</td>
<td>0.222</td>
<td>0.120</td>
<td>0.823</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td>0.252</td>
<td>0.571</td>
<td>0.906</td>
<td>0.930</td>
<td>0.942</td>
<td>0.563</td>
</tr>
<tr>
<td><strong>Reading Level</strong></td>
<td>0.536</td>
<td>0.421</td>
<td>0.623</td>
<td>0.477</td>
<td>0.562</td>
<td>0.554</td>
</tr>
<tr>
<td><strong>Dyslexic</strong></td>
<td>0.174</td>
<td>0.440</td>
<td>0.443</td>
<td>0.681</td>
<td>0.336</td>
<td>0.512</td>
</tr>
</tbody>
</table>

As the above chart shows, none of the demographic or reading-skill measures appears to have a statistically provable influence on any of the app response scores. All of the associations fall well short of the commonly accepted measure of certainty at $p <= 0.05$. 

---

**Figure 22 – Experiment 1 Combined Ratings Scores**

**Figure 23 – Experiment 1 Data P Values**
Even at the more generous limit of $p \leq 0.10$, the data cannot be taken as proof of a relationship, although two measures fall close at 0.120 and 0.138. These relate respectively to a possible link between age and likelihood of using the app in future, and age and considering the app better than paper books.

Graphing the first of these two possible relationships, it does appear that there is a difference in the percentage of respondents in each group to give each rating, but an overall trend is unclear. The distribution patterns suggest that the 8-year-olds were slightly more likely to want to use an app like this in future, whilst the 9-year olds were almost evenly split between Yes and No, with the sharpest opinion divide. The age 10-11 grouping was also split, and overall had the highest incidence of #5 responses. This may, however, be influenced by the fact that the age 10-11 group was smaller than the others, therefore a single response carried more weight.

At present it is difficult to say whether there is any meaningful relationship between age and wanting to use the app in future, and further data from a wider and more balanced range of age groups would be necessary to make any concrete statements.

Charting the second potential relationship, age vs. preference for app over paper, the picture is far clearer. Here it is possible to see a strong difference in the preference of the 8-year-old group as compared to the two older groups. The 8-year-olds appear far more likely to prefer the app strongly over a paper...
book, whilst opinion was much more mixed amongst the 9-11 age groups.

This is a promising link which may warrant more research in order to determine whether, with a larger sample of age groups, the app format might be statistically proven as more enjoyable than traditional paper books for some ages.

The only other potential relationship which might be worth looking into would be a comparison of enjoyment vs dyslexia diagnosis status. At $p = 0.174$, and with only two confirmed dyslexic participants, it is too uncertain to draw meaningful results on this from this study. However, further testing with a larger population of dyslexic users would determine whether the app is in fact more enjoyable to users with the condition.
6.3 – Results – Comment Feedback

The second half of the data from this experiment comes from the three open-comment boxes at the end of the questionnaire, in which subjects were invited to leave any positive, negative or miscellaneous (including suggestions) feedback.

This data was analysed by manually compiling a list of keywords extracted from each response, staying as true to the original statements as possible with occasional minor rephrasing for clarity/spelling. Some feedback was also moved into a more appropriate category, i.e. suggestions in the negative feedback box were shifted into the miscellaneous list and vice versa.

This produced three lists of key words and phrases which were then analysed for frequency of occurrences to see whether any trends emerged. Some subjects left several comments per box; others left none, therefore all keyword counts are shown as a percentage of comments made rather than comments per participant.

To begin with an at-a-glance view of which terms subjects mentioned most frequently, the three lists are represented below in word-cloud format. For the purposes of these charts, synonyms (such as options and settings, video and videos) were merged into a single term to show frequency more accurately. Each feedback set is shown with words sized by frequency at 100% weighting:

(Full-size copies of each of the word-cloud images are available in the Appendices).

![Figure 26 – Positive Feedback Keywords Cloud](Image)
As illustrated above, there were some clear trends in the written feedback, with several key terms appearing repeatedly in each section.

Below is a categorised breakdown of response type and frequency for each feedback set:
In terms of positive feedback, by far the most mentioned aspects were the multimedia elements; the images, videos, sound effects and UI graphics (examples from the test app shown below: from left, title screen background, customised text slide screenshot, illustration image).

Counting all references to multimedia content, this category accounted for 96 of 193 total positive keyword/phrase entries, or 49.7%. This overwhelming majority shows just how much of an impact the use of multimedia had in attitudes towards the app. This corresponds with the results of previous research which found that the inclusion of multimedia content in eBooks can increase enjoyment and engagement in children (Takacs, Swart and Bus, 2015).

The second most frequently mentioned topics, tied at 28 occurrences each (14.5%), were general comments on the app being enjoyable/engaging and the progression mechanic. General comments included 3 specific mentions of the app being more enjoyable than paper books, plus one comment from a user who “didn’t like reading but liked this app.” This backs up the Likert score results which showed that for at least some participants, there was a marked improvement in enjoyment over a typical paper book.

As for comments on the app structure, the ability to choose your own path and unlocking new levels were mentioned most often in this category. This suggests that the use of the branching unlock structure was a reasonable success.
Third most popular by comment frequency were elements of the stories themselves, with most mentions going to humorous content such as a video with evil-witch-laugh sound effects (which members of one class spent much of their session imitating!) and a chapter featuring Red meeting a group of frogs obsessed with getting someone to kiss them (illustration pictured right). This relates to the findings of (Scholastic Inc. and YouGov, 2015), in which researchers found that amongst this study’s target age group, humour was the most popular genre of reading content.

The remaining two categories of response received concerned customisation (9 comments) and difficulty (4). In customisation, respondents particularly liked being able to set their own font options, with mentions made of text size, font and colour.

Difficulty was least mentioned, with two comments saying that users liked the range/difficulty of vocabulary used. Another liked the ratio of words to pictures favouring text. Only one mention was made of the ranging difficulty levels, with one response saying that it would be good for users of all ages.

This is a disappointing result given that the variable difficulty was intended as a key draw for users, but this factor may be down to time constraints during testing. Due to the time available in the session, many of the participants only had time to complete the content once on the default medium setting, making comparing the different levels difficult. Further testing over a longer session would yield more data on whether variable difficulty might successfully influence reading positivity.

Finally, miscellaneous positive comments mentioned saving trees, kid friendly design, the name of the app, and having their class featured in the credits for any eventual app release.
• **Negatives**

Regarding negative comments, there were a few dominant themes in what users disliked about the app.

The “winner”, with 25 of 99 mentions, was app stability. Due to a bug which was thought to have been fixed before user testing began, the app would experience a terminal error in certain circumstances and crash. As the patch put in place beforehand proved ineffective, this bug caused some problems, especially as at this stage, the app was clearing user progress on close. This had been intended as a time-saving feature to allow users to complete a session, close the app and pass the tablet directly to the next test subject without having to manually clear data. This combination resulted in frustration for users, many of whom who lost their progress mid-session and had to restart.

It is therefore unsurprising that glitches and restarting were amongst the most frequently used terms in the negative comments, and highlighted the need for more time spent double-checking bug testing before giving an app to users. This was taken on board for development of the second prototype, and stability fixes were one of the areas prioritised for improvement.

Second most popular amongst the negative comment categories, at 18 comments, was length/amount of content – specifically, the phrase “not enough.” The most common complaint was that there was only one book available rather than a selection of different ones. This complaint may be taken as something of a back-handed compliment to the app – rather than lacking interest, users wanted *more* content to play through.

However, not much could have been done to address this particular complaint with the time and resources that were available. Creating a complete, if short, book worth of content including illustrations, with variations for three difficulty levels, already took a significant amount of time from software development. As the functionality and effectiveness of the app as an example of a multi-level eBook system was at the core of the project, rather than the actual content used in testing this, the volume of content produced had to be balanced against other concerns.

The third most frequent category of criticism regarded difficulty, at 12 comments. The majority of these complaints were identical: “too easy,” or words to that effect. From comments made both during the testing session and in the survey responses, it became quite clear that the choice of a Grade 3 (age 8-9) difficulty level as the middle-difficulty benchmark was simply not challenging enough for the majority of test users, even in the younger classes. Several complaints made mention that they found the content was aimed too young for them. Only one comment about difficulty ran in the opposite direction – one user had found it too tricky at parts.
This result suggests that getting an accurate measure of users’ reading levels before designing the text content would have been a more appropriate methodology than simply basing the difficulty on an age-based benchmark. Whilst it was not possible to administer a reading level test to participants during the initial design stage in this case, future work should be carried out with this possibility in mind.

Users’ comments on their experience with the difficulty during the test session also brought up another point to be improved: just about every user jumped directly in to start reading without looking at any of the settings, and therefore ended up using the default medium difficulty. This meant many users who might have been satisfied playing the hard mode content ended up with material too simple for them.

This highlighted a design flaw in the app flow, which was corrected for Prototype 2 by adding an initial difficulty-selection screen which would appear on clicking the “read” button for first-time users before they could progress to the content.

Next amongst the negative comment themes was that of media and graphics (10 comments). The most common complaint here, mentioned by 4 of the 10, was the creepy-laugh witch video which also featured in several of the positive comments (6 mentions). This suggests that whilst this content was appropriate for the majority of users, and indeed was enjoyed immensely by some, it may have been a little on the scary side for some of the most sensitive children. Notably most of the complaints about this content came from the youngest users, so a difference in maturity levels may be at the root of this split.

Whilst this content was well within the boundaries of what would be shown in a PG movie, from this feedback it can be concluded that the graphic should have probably been swapped out for something more humorous and the original sound effect left in.

Other visual aspects the users disliked were the style of some of the videos and graphics, plus the font selection, which one subject labelled “boring”. Given the relatively small number of negative comments in this area compared to the volume of positive responses, this may simply be down to personal taste and the fact that no single graphical style will be considered attractive by everyone.
The remaining specific comment types, at 9, 8 and 4 responses respectively, were content, structure and enjoyment/engagement.

In terms of content, users disliked that the story used was familiar and unoriginal (although one also disliked the fact the story was changed from the original!). This could be expected considering that it was based on a common fairy tale.

Regarding structure, some subjects disliked the way that the story was set out, whilst others were not fond of having to unlock content. For the latter, the inclusion of a freestyle, everything-unlocked mode as considered in the early stages of the project might improve their experience. Last amongst the complaint categories was enjoyment and engagement, in which four users stated that they found it boring or not very much fun. Whilst disappointing, this represents only a small proportion of users.

There were also 13 miscellaneous negative comments, including a few general mentions along the lines of “I didn’t like it” and some more constructive comments, such as reading on a tablet being bad for the eyes at night (this user may not have tried Dark Mode, which was designed for this), encouraging kids to spend more time on mobile devices, and potentially being a threat to traditional authors.

The lack of a clear indication of which node had just been unlocked was also mentioned, and so a glowing effect was added to the buttons for newly-unlocked content in the second prototype.

Taken collectively, the negative comments produced a useful list of areas to be fixed or improved for further versions of the app. These included prioritising stability fixes and bug testing, creating future content with a better idea of appropriate difficulty levels and some restructuring suggestions.
In a theme continuing on from the negative comments, by far the most frequently occurring single word in the miscellaneous comments/suggestions section was “more,” with 38 occurrences in 59 comments.

Of these, the most requested area in which the users wanted “more” was content – 23 of the 59 (38.9%) comments were asking for more books, levels or stories. As mentioned previously, this can be taken as a measure of the app’s success; the fact that test subjects were in fact asking for more demonstrates a level of enjoyment and engagement in line with the positive Likert scores.

There were also 8 requests for more multimedia content, an unsurprising result given that this was the most frequently cited aspect of the app when users were asked what they liked.

The second most frequent category of response amongst the miscellaneous comments was additional feature requests. These included:

- Audio function to read out the story (2 requests)
- Minigames/word games (2 requests)
- Freestyle reading mode
- Background images for text view screen
- Books specifically for kids with learning disabilities
- Captioned video views to read
- Auto-play mode showing content in a given order

Amendments to existing features were also requested:

- Put difficulty choice at the start (implemented)
- More vocabulary
- More difficulty levels
- 3 requests to remove the creepy witch video if using the app with younger kids (from those who commented on this previously.)

There were also 9 requests for different types of stories, asking for older-aimed stories, different genres and more original storylines. One user also suggested using the app for non-fiction books.

Overall the miscellaneous comments were filled with positive suggestions, some of which match up with potential further development goals already suggested early in the project. This shows that building these additional features into any future version of the app would be worthwhile, and gives an indication of which to begin with: audio content and game elements.
7 – Experiment 2

7.1 – Implementation

Experiment 2 was conducted in early August to gain feedback from a group of young adult (age 18-30) users, using an improved Beta build of the app. A screenshot showing part of the redesigned map screen is shown right in Figure [32].

As mentioned previously in Section 4, this experiment was affected considerably by scheduling issues; namely that it was not possible to find a day during the available time period in which all seven of the participants who had agreed to take part could attend a meeting. The format of this experiment was therefore changed from the initial plan of conducting a live focus group session to allowing participants to submit their feedback individually via an online form. This decision was chosen over recruiting for more test participants as a smaller amount of guaranteed feedback was considered better than the uncertainty of recruiting a new test group in a short time period.

This solution was not ideal, however, and came with its own problems as it relied on participants being able to test solely on their own devices rather than being provided with hardware to use. Of the seven participants, five were able to successfully test the app and upload their feedback, whilst two only had access to older Android devices which were not able to run the software.

Subjects were asked to take the time to play through the app, with a minimum requirement of reaching at least one ending and trying more than one difficulty mode. They were then asked to post their feedback via a series of open-answer questions.

Although this experiment yielded a disappointingly small amount of data compared to the original planned session, it can still provide some insight into how the young adult demographic viewed the app, and gives some idea of areas to target for future work.
7.2 – Results

Participants were asked 9 open questions about their reading habits and their experiences using the app. Some questions were multi-clause due to question number limits imposed by the survey service being used. There were six specific questions, followed by three general comment boxes which were identical to those used in Experiment 1.

A summary of the responses to each question is given below:

- **Q1 – Please describe your current habits as regards reading for leisure (how often, what kind of materials)**

There was a good range of reading habits amongst the participants, with three frequent readers (daily or almost daily), one infrequent reader and one who stated that they did not read for leisure at all. One of the frequent readers also mentioned that their primary method of reading was via a Kindle device, meaning that there was at least one experienced eBook user in the group.

This group represents a reasonable cross-section of different reading habits, although it is somewhat weighted in favour of those who already read for leisure given that 3 of the 5 were already frequent readers. Compared to figures on young adult reading levels (Huang et al., 2014), (Mokhtari, Reichard and Gardner, 2009) already examined, this seems a high proportion, although it should be considered that whilst the quoted studies used college students for their test subjects, this group were from a range of academic backgrounds and occupations. This may be a potential factor affecting their reading habits.

Whatever the cause of the difference, the three frequent readers’ responses will be considered as likely to have a positive skew towards reading and will be treated appropriately.

Of those who gave information on their preferred type of reading, sci-fi and fantasy fiction were most popular, with one subject also liking factual reading about technology.

- **Q2 – Overall, what was your first impression of the app?**

The responses to question 2 were positive, with users focusing largely on ease of use and UI style. The app was described in terms such as “welcoming,” and “uncluttered.” The fact that ease of use was mentioned more than once suggests that this factor was particularly important in making a first impression on users.

This would suggest that making the UI as simple as possible, based on the minimal design of the Android Material Design style, worked well as users were able to navigate the app easily.
- **Q3 – Did you find using the app enjoyable and/or engaging?**

All five users answered this question in the positive, which was an excellent result for the app. The path choice aspect was stated by three users as their reason for finding it enjoyable, in addition to one mention of enjoying the multimedia aspect.

This paints a fairly clear picture that the branching structure and the way that users can customise the story through their own choices were a hit amongst this group, suggesting that this aspect may be an area to focus on with young adult users.

- **Q4 – How did using the app compare to an ordinary paper book?**

Regarding this question, four of the five users indicated a clear preference for the app format as opposed to a paper book. One user described it as “more absorbing;” another called it “much more exciting.” There was also a comment comparing the experience to reading a traditional “choose your own adventure” novel, saying that this format offered the chance to make choices in a more streamlined, user-friendly way.

The user who favoured traditional books said this was down to the tactile experience of reading on paper, something an app cannot replicate. They did however state that they thought the app would be “great for reading on the go”.

- **Q5 – Would you read books in this format if they were available, and if yes, would you read more often using the app than you read currently?**

All of the test subjects responded to this question with a yes to the former, and three of them said the same to the latter.

Interestingly, of the three, only one was already a frequent reader. In fact, the most definite responses stating that users would read more often using the app came from the infrequent and non-readers.

This suggests that the app would have more of an effect on the habits of infrequent readers than those who already enjoy reading regularly. As a tool to encourage reading amongst reluctant readers, the positive response from users in that demographic is an encouraging result.

- **Q6 – What did you think about the structure of the content (in particular the branching choice paths and multiple difficulty modes)?**
Again this question received positive responses from all five users. The area most commented on was the branching path system, with three users citing this as something they particularly liked.

Users also liked that the customisable options made the app suitable for readers of all ages.

Only one specific mention was made of the difficulty system, similarly to the feedback from Experiment 1. This is again disappointing, but there were other comments about how the app could be “tailored to the user’s needs” and being “in control of how the story was told,” which may also be considered to refer in part to the difficulty settings. This at least suggests that the difficulty system was positively received, even if it was not considered as interesting a factor as the branching choice system.

- **Q7 – Please comment about anything you particularly liked about the app:**

The choose-your-path mechanic and the use of multimedia content were tied as most popular in the responses to this question, with three mentions each.

This corresponds largely with the results from Experiment 1, in which these two aspects were also the most cited in the positive comments.

There were also two comments regarding the map layout used to represent the story timeline, suggesting this form of interface was an effective way to show the path structure.

- **Q8 – Please comment on anything you particularly disliked about the app:**

The responses to this question also echoed those from Experiment 1 – the only topic of complaint here was that certain areas of the app were prone to lag or crashing on their devices. From users’ descriptions, this appears to have been largely down to memory issues on older or less powerful devices. Whilst the app should support android devices back to API 11, the media files used are relatively large, and this result suggests that a trimmed-down version for older devices, particularly phones, may be necessary to properly support these devices.

- **Q9 – If you have any additional comments or suggestions, please enter them here:**

Only three miscellaneous comments were left; one was a general statement that the user had enjoyed the app and would like to see more books added.

The other two were suggestions; one content-oriented (more adult-aimed books, with more cut-throat choice options) and one regarding the interface. That suggestion was for a
separate “start” button which would take the user through a first-time setup of the app rather than including it under “read” on the first run. This could easily be added as an option for future iterations and should be considered as a way to make the new-user experience even simpler.
8 – Discussion

In order to assess the overall success of this project, several questions must be considered:

- Was the project well implemented?
- Were the research questions answered?
- Were the results meaningful and useful?
- What further work could be done to improve on the results?

Each of these topics will be discussed in this section.

8.1 – Project Implementation

Whilst it was completed successfully, this project was not without its share of problems throughout.

Regarding Experiment 1, there were several issues which arose. First amongst these was the challenge of arranging the collaboration with a school in order to conduct the test. When the idea of testing in schools was first brought up, it was initially dismissed as too difficult to arrange in the time period available, but as it emerged that this would provide the best data to assess the project, an approach was made.

There were potentially costly delays of several weeks in communication with the school staff, but finally a response was received stating that Echline Primary would be willing to host the testing session. Thankfully the process of seeking a host school was begun very early in the project, so the delays in getting a response were not as damaging to the project schedule as they may have been if left later.

Following this, the process of getting parental permission for the children to take part was begun. An explanatory letter and permission form was written and given to the school for distribution. Again, it was fortunate that this was conducted more than 6 weeks before the session was due to take place, as the initial response rate for signed slips returned was extremely poor. Eventually appropriate permissions were secured with some help from the school and the session was able to go ahead.

The session itself was hit by technical issues – namely the pre-loaded “offline” survey app which was intended to be used for collecting feedback turned out to require its online login to be re-entered any time someone closed the app. This was in part a failure on the researcher’s part to test the survey app thoroughly enough before use. An attempt was made to connect to the school’s WiFi network to fix this, but authentication issues occurred. Not wanting to waste any further session time trying to fix the network issues, it was decided to simply let the classes test the app and come back with printed feedback sheets
to distribute, as per the contingency plan set out during risk analysis. Whilst not ideal, having backup plans in place allowed the session to proceed with far less disruption than might otherwise have been experienced.

Waiting for the paper surveys to be returned and then manually digitising the data did cause some further delays, but overall the experiment did achieve what it set out to do; a large data set concerning reactions to the app was successfully collected. Experiment 2 also had its issues, primarily the common research problem of recruiting enough participants to get meaningful data. In this case, a group of seven were initially signed up, but due to conflicting work schedules it was not possible to get them all to a group session. Once again a contingency plan had to be implemented and the subjects were allowed to test the app on their own and submit written feedback online. Two had to drop out due to not having compatible devices, but detailed feedback was still collected from the five others.

Overall, although both experiments hit organisational difficulties, thanks to a solid risk analysis at the beginning of the project, preparations and contingency plans were put in place for as many potential problems as possible. This allowed the numerous difficulties experienced along the way to be overcome and eventually resulted in the collection of two strong data sets. The flexible development schedule was also invaluable here; development blocks were shifted and rearranged as needed to fit around delays whilst still keeping the project on track.

8.2 – Questions and Answers

A successful experiment is no good without meaningful results, and so the next question to be asked is, did the project manage to answer its research questions?

Rephrasing the hypotheses defined in section 3, the research questions for this project were:

- Is an interactive storybook app more enjoyable than a traditional book/eBook?
- Is an interactive storybook app more engaging than a traditional book/eBook?
- Is an interactive storybook app more motivating in terms of reading more often than a traditional book/eBook?
- Would users who find reading difficult/uninteresting read more often using an interactive storybook app?

Given the combined results from Experiments 1 and 2, it may be said that on all counts, these questions were answered to a reasonable degree.
A moderate positive trend in general enjoyment was shown in rating data from Experiment 1, which the open comments from both experiments also supported. In general, the majority of users seemed to like using the app, and many showed a significant level of enthusiasm. Whether this enjoyment factor could be sustained over a longer period without the advantage of novelty remains to be seen, but the data is promising. The areas which contributed most to enjoyment were found to be multimedia content and the choose-your-path system, showing that interactivity can enhance the eBook experience.

When asked whether they preferred using the app over a paper book, again a majority responded in the positive. Data from Experiment 2 suggests that this preference may be strongest in those who were infrequent leisure readers, but the Experiment 1 results do not have enough data on this topic to confirm. This may have been a missed opportunity as the Experiment 1 survey did not ask subjects how often they read for fun, focusing more on reading ability. This gap would be something to address in any further testing.

Overall the data from these questions suggests that in a reasonable majority of cases, the app was found more enjoyable than a paper book, but the paper novel still does have its fans who are yet to be won over.

In terms of engagement, rating results from Experiment 1 were middling, although a number of positive comments along the lines of “it held my interest” or “more interesting” were received. Experiment 2 got more definitive answers on this topic, with a strong indication that users did find the app engaging, particularly due to the aspect of choice and control. In this case the key question of “is the app more engaging” can still be answered in the positive, although the response in this case would be more tentative and in need of further data for confirmation.

Regarding the question of whether the app could successfully get users – especially those who were struggling or infrequent readers – to read more often, which is arguably the most important of the research questions, results were promising.

In Experiment 1, a majority of participants answered in the positive regarding both whether they would use the app in future and whether it would encourage them to read more often. One comment in particular was the exact response that was hoped for from reluctant readers: “I don’t like reading but I like this app.” Experiment 2 participants gave similar responses, with the infrequent readers especially saying that they would want to read more often given books in this format. Those who were already daily readers were less convinced that it would change their reading habits, but as the idea behind this app was to provide a way to engage those who do not already read much, the result still shows that the app achieved its goals.
8.3 – Results Summary

To sum, all of the results from this project provide a solid though not definitive answer to the questions that it set out to answer. There is enough data to support all four of the original hypotheses, although not enough yet to call them proven.

One of the limitations of this study is that all of the data collected is very “soft”: even the quantitative data from the Likert scales is based on subjective, opinion-based answers. In order to really prove the research hypotheses, hard, quantifiable data would be needed.

The data also comes from only one brief test session with the app; enough for a first impression, but perhaps not enough to determine its potential for long-term use.

The app was also relatively simple – it met all of the red “must-have” requirements defined in the software checklist, plus it included some of the optional content, such as hidden nodes, but more could be added to improve the experience in future versions.

Nevertheless, the overall trend towards successfully encouraging people to read more is positive enough that this area would be promising for future study – a number of suggestions for which are detailed below.

8.4 – Future Work

The next stage for this project would be to develop the app further, including a larger range of content and some of the additional features suggested by users.

This could then be used for a longer-term experiment in order to collect hard performance data. A good format for this would be another school study, this one of of at least 8 weeks, in which a control group and an app user group had their reading scores and time spent reading measured over an extended period. A within-groups model would be most suitable for this, with participants swapping roles halfway through, to account for user variation in ability, although it would mean a longer test period.

This would give solid data on not only whether using the app encourages users to read more often, but also allow researchers to define just how much more often, and how much of an effect this may have on improving reading ability over time.

Another factor which needs more testing is the effect of the variable difficulty system on user enjoyment/engagement, and this could be tested by creating better text content for future tests, benchmarked using the subjects’ actual reading levels. This factor may in fact be suited to being tested in its own experiment, as above but with one user group using a non-variable difficulty version of the app and another given the level choice.
Further testing with a wider range of age groups would also help answer the question of whether this type of app is more effective with certain ages than others.

In terms of additional development of the app itself, other than adding more content features, the next important tool to develop would be to split the book creation process off into a separate app with its own GUI rather than creating content nodes via code, so that book content could be created more quickly and easily.
9 – Conclusions

In an area as widely researched as reading and storytelling, it is often difficult to see where to begin to find out something new. When this is combined with creating content for the ever-changing and constantly developing world of mobile software, this problem appears to multiply exponentially.

This project aimed to meet this challenge by testing a combination not analysed before in eBooks: a game-like pairing of variable difficulty settings with a choose-your-own-adventure style branching path system.

The focus of testing was whether this combination of user-driven choices with the multimedia content of an eBook could encourage users to read more often.

After conducting a series of experiments with two different user groups, collecting feedback from a total of 111 participants, answers to this question began to emerge. Whilst opinion varied greatly across the range of subjects, there was an overall positive trend towards the eBook app. The majority of users who tried *Wander* responded that they had enjoyed using the app and would want to read more in future if an app like it were available.

Whilst the results were not enough to state that multi-level eBook apps are the perfect solution to all literacy problems for everyone everywhere, there is some promising data to suggest that for at least some users in the age groups tested, this type of app can indeed help improve time spent reading and thus aid literacy scores.

This project has been a huge challenge to develop and test, and to achieve such a positive result is the best outcome that could have been hoped for.

It is hoped that from here this concept can be developed even further. Like any good research project, this one has discovered both answers and new questions, and there is much that could be done from here in terms of both further development of the app and deeper research into the effects of this software on readers.

In terms of ultimate goals for this software, to make it possible for authors to create and publish complete original works using the *Wander* system would be the highest goal to aim for. Seeing this format come into real-world use so that a broad variety of users could enjoy and benefit from such a system would be the project’s ultimate achievement.
10 – References


Conboy, G. et al., 2011. EPU 3 Overview. [online]


Dyslexiefont.com, 2015. History of the. [online]


http://crl.acrl.or


