Institutionen för teknik och naturvetenskap
Department of Science and Technology

Examensarbete

Developing a Mobile Learning Application for Entrepreneurship Education in Uganda and Zambia

Examensarbete utfört i Medieteknik
vid Tekniska högskolan vid Linköpings universitet
av

Marcus Nygren

LiTH-ITN-EX--16/NNNN--SE

Norrköping 2016

Department of Science and Technology
Linköpings universitet
SE-601 74 Norrköping, Sweden

Linköpings tekniska högskola
Linköpings universitet, Campus Norrköping
601 74 Norrköping

Developing a Mobile Learning Application for Entrepreneurship Education in Uganda and Zambia

Examensarbete utfört i Medieteknik
vid Tekniska högskolan vid Linköpings universitet
av

Marcus Nygren

LiTH-ITN-EX--16/NNNN--SE

Handledare: **Lena Tibell**
ITN, Linköpings universitet
Iliana Björling
YoungDrive

Examinator: **Camilla Forsell**
ITN, Linköpings universitet

Norrköping, 11 augusti 2016

Avdelning, Institution
Division, Department

Media and Information Technology
Department of Science and Technology
SE-601 74 Norrköping

Datum
Date

2016-08-11

Språk
Language

☐ Svenska/Swedish
☒ Engelska/English
☐ _____

Rapporttyp
Report category

☐ Licentiatavhandling
☒ Examensarbete
☐ C-uppsats
☐ D-uppsats
☐ Övrig rapport
☐ _____

ISBN
__

ISRN
LiTH-ITN-EX--16/NNNN--SE

Serietitel och serienummer
Title of series, numbering

ISSN
__

URL för elektronisk version
http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-XXXXX

Titel
Title

Developing a Mobile Learning Application for Entrepreneurship Education in Uganda and Zambia

Författare
Author

Marcus Nygren

Sammanfattning
Abstract

Entrepreneurship educations in developing countries have not yet been able to take advantage of digital tools. The Ugandian non-profit YoungDrive has 60 coaches teaching entrepreneurship to 12 000 youth in rural areas. The coaches have a problem during and after their education with assessing and improving their abilities to learn and teach entrepreneurship. The purpose of this study was to investigate how an app can be designed to address this issue.

Methods within service design, agile development and interaction design has been used and combined to construct and analyse interviews, workshops, question sets, and app tests with the coaches in Uganda and Zambia. In total, three months were spent testing and iterating on low-detailed and high-detailed prototypes. The result is a launched hybrid app for Android, iOS and web.

A formative test shows coaches are more reliably correct using an improved design of multiple-choice questions than a standard multiple-choice design. Interviews shows the coaches has become more aware of what they know and do not know, and feels more confidence before their youth lesson with an increased quiz result. Further research should evaluate that the actual quality of the youth lesson improves.

Increasingly well-constructed multiple-choice questions with thoughtful feedback could stimulate creativity and problem-solving, deemed important by entrepreneurship education research. After overcoming usability issues, the final app could reach both low and high-order learning objectives within entrepreneurship.

The app did seemingly improve the quality of entrepreneurship education for the coaches in this specific developing world context. Further research should also investigate the design and implications of a digital-only entrepreneurship education for the coaches, having in mind that the teacher is believed the main factor of entrepreneurship education. As of now, the app is an effective compliment and assistance to the physical training.

Nyckelord
Keywords

entrepreneurship education, digital learning, mobile app, developing countries

Abstract

Entrepreneurship educations in developing countries have not yet been able to take advantage of digital tools. The Ugandian non-profit YoungDrive has 60 coaches teaching entrepreneurship to 12 000 youth in rural areas. The coaches have a problem during and after their education with assessing and improving their abilities to learn and teach entrepreneurship. The purpose of this study was to investigate how an app can be designed to address this issue.

Methods within service design, agile development and interaction design has been used and combined to construct and analyse interviews, workshops, question sets, and app tests with the coaches in Uganda and Zambia. In total, three months were spent testing and iterating on low-detailed and high-detailed prototypes. The result is a launched hybrid app for Android, iOS and web.

A formative test shows coaches are more reliably correct using an improved design of multiple-choice questions than a standard multiple-choice design. Interviews shows the coaches has become more aware of what they know and do not know, and feels more confidence before their youth lesson with an increased quiz result. Further research should evaluate that the actual quality of the youth lesson improves.

Increasingly well-constructed multiple-choice questions with thoughtful feedback could stimulate creativity and problem-solving, deemed important by entrepreneurship education research. After overcoming usability issues, the final app could reach both low and high-order learning objectives within entrepreneurship.

The app did seemingly improve the quality of entrepreneurship education for the coaches in this specific developing world context. Further research should also investigate the design and implications of a digital-only entrepreneurship education for the coaches, having in mind that the teacher is believed the main factor of entrepreneurship education. As of now, the app is an effective compliment and assistance to the physical training.

Acknowledgments

Due to a chain of lucky events, this master thesis took the approach of combining service design, thoughtful interaction design, technology, learning effectiveness research, and entrepreneurship.

For service design, I want to thank Peter Gahnström at LiU Innovation leading me to Expedition Mondial. There I especially want to thank Susanna Nissar for being a great tutor, and also Erik Widmark.

Thoughtful interaction design was introduced thanks to a recommendation from Lena Tibell and Konrad Schönborn, to go to Jonas Löwgren's first test lecture about interaction design at Linköping University, which led me to looking deeper into his literature, and discovering the book Thoughtful Interaction Design, which was a perfect fit into the world of interaction design for me, coming from an engineering perspective. It showed me what being a great designer really meant, and I was compelled.

For learning effectiveness research, I greatly want to thank Lena Tibell and Konrad Schönborn. Henrik Marklund at Knowly has also provided a valuable perspective on digital learning.

For entrepreneurship education, I have so many to thank, including again Konrad Schönborn. But by biggest thanks goes to those that knew the entrepreneurship coaches: Josefina Lönn and Iliana Björling from YoungDrive, and Gerald Emoyo from Plan International. Also, Julien Tantege from Grameen Foundation in Kampala shared her experience working with technology targeting the very same coaches.

Thank you for your contributions! Lastly but not least, I want to thank Linnea Rothin who put me right track on the master thesis, and to Lena Tibell and Konrad Schönborn who has been supervising this master thesis from start to end. Your expertise and continuous feedback have been incredibly valuable, and I feel honoured to have been given this learning opportunity.

Thank you.

Norrköping, August 2016
Marcus Nygren

Contents

1

Introduction

The following master thesis was carried out in the Master's program of Media Technology and Engineering at the Department of Science and Technology, Linköping University. The work has been carried out for the non-profit entrepreneurship academy YoungDrive in Uganda and Zambia. The organization teaches starting a business to youth in developing countries, with YoungDrive coaches now educating 12 000 youth within Plan International's project in Uganda, and now also running a coach training in Zambia. An application for the entrepreneurship coaches is thought to help coaches assess what they know, and learn to become even better with teaching entrepreneurship to the youth.

There are both opportunities and challenges for mobile tools for rural areas of Uganda. While entrepreneurship activity is high, and mobile services are growing rapidly (see 3.1.4), the two fields have not yet been combined with entrepreneurial learning in a developing country context. To properly answer YoungDrive's need, entrepreneurship education and digital education has been researched together with how to develop apps for web and mobile platforms simultaneously. The research questions focus on the process and result of developing an app used by first-time smartphone users for entrepreneurship learning.

1.1 Purpose

In order for young ambitious entrepreneurs to build sustainable enterprises, they need to have basic entrepreneurial skills. A mobile learning platform could assess and teach those skills either complementing a physical entrepreneurship education, or being accessible to the entrepreneur whenever and wherever deemed most important.

1.1.1 An App for the Entrepreneurship Coaches

The entrepreneurship education YoungDrive is an initiative of Illiana Björling from YoungDrive, now collaborating with Plan International. Within the project "A working future", they have educated, supported and inspired 12 000 Ugandan youth in the process of starting their own businesses (Nissar, 2016).

The overall aim of the master thesis is to develop a working prototype of an entrepreneurship coach training app. The master thesis is about how to design an app for entrepreneurship education, including evaluating it's effectiveness towards the coaches. The app should be suited for the coach training, and have multiple-choice questions as its basis mechanics.

1.1.2 Definition of Success for the App

By training coaches that can carry out the education in larger groups of entrepreneurs, the education reach many young people at the same time. A mobile learning platform is predicted to improve the effect of the training even more, by fulfilling the following purposes:

- Validate the coaches' level of knowledge during their education
- Train the coaches on distance
- Certify all staff

YoungDrive's experience goal for the app is "It should be easy to understand, pedagogical and enjoyable to use, and the coaches should think it is fun and meaningful to learn via the app" (Björling and Lönn, 2016).

1.2 Research Questions

The overall aim of the study is to create and apply a design process of an application for entrepreneurial learning, to be implemented in a developing country context. In response, the following specific research questions were raised:

1. How is the development affected by the technical possibilities?
2. How is the design affected by the contextual constraints?
3. How can test questions be developed to support entrepreneurship learning?
4. How does design affect usability and learning done via the app?
5. How can users' feedback be used to inform modifications of the app?

2

Theory

To understand how to reach the objectives of the project, this chapter describes and presents related work on entrepreneurship education and digital learning. An effective learning app must help the user to reach educational objectives, which is why methods to design for learning is presented. Further, to ensure a learning app will be used voluntarily, methods to design for motivation is presented from a product development standpoint.

2.1 Entrepreneurship Education

Entrepreneurship education has been a growing field of investigation over the last three decades. While Dickson and Weaver (2008) says there are few empirical studies available, examples include among others Kuratko (2005), Pittaway and Cope (2007) and Bae and Fiet (2014). Oviawe (2010) and Iakovleva and Stehphan (2011) has had an interest in interventions for teaching and learning entrepreneurship in the developing world.

First, Oviawe (2010) conclude by how teaching of creativity and problem-solving skills seems to be especially beneficial for entrepreneurship in developed countries. In YoungDrive, the youth are tasked with starting their own business from no capital, which fosters creativity and problem-solving skills. Further, Iakovleva and Stehphan (2011) indicated that respondents from developing countries do have stronger entrepreneurial intentions than those from developed countries. This stems from attitudes, subjective norms, and perceived behavioural control. Their encouragement is that developing countries need to focus on the development of institutions that can support entrepreneurial efforts. YoungDrive is one such example.

Ruskovaara and Pihkala (2015) concludes, that the teacher seems to be the main factor for entrepreneurship education, and that research agrees with them.

There seems to be no indication of difference between men and women, nor previous professional teaching experience. They could find that entrepreneurial activity seems to lead to better entrepreneurship education. Dickson and Weaver (2008) recommends mainly two things for enhancing entrepreneurship education practices. First, the playful side of teaching and learning is mentioned. Secondly, they encourage teacher training that develops the competencies as a mentor, enabler or coach.

2.2 Learning with Mobile Technology

YoungDrive has asked for a multiple-choice question learning game used for assessment and improvement of the YoungDrive learning objectives during and after the training. In recent times, digital learning (e-learning) has had a tremendous impact both outside and inside the classroom. With a growing teacher interest, research so far shows that digital education is difficult and risky, but potentially rewarding (Luckin and Noss, 2012). Thus, digital education simultaneously shows great potential and vulnerability. Much of the research on digital learning to date on digital games has focused on proof-of-concept studies and media comparisons (Luckin and Noss, 2012). A study by Clark and Killingsworth (2014) motivates why a digital tool or game is a good thing by showing an increase in learning outcomes, relative to non-game instructional conditions.

A large amount of development has taken place on diagnostic testing environments, that allow teachers and learners to assess present performance against prior performance (Luckin and Noss, 2012). There are numerous examples of developments in *e-assessment* using mobile environments, as well as immersive environments and social and collaborative environments. One such example is the educational app platform iSchool, developed by iSchool Zambia (iSchool, 2014). The app has been praised and made popular as it was designed to fit the Zambia school curriculum to the point, accessible as a home edition, pupil edition and teacher edition. Interest in formative e-assessment is increasing. It has been shown that multiple-choice tests in e-assessment can be used to good effect (Nicol, 2007).

Two studies within electronic assessment (e-assessment) or mobile learning (m-learning) have been done that this master thesis is inspired by. One uses deliberate practices in a mobile learning environment (Yengin and Uzanboylu, 2012). The other focused on and further validated the research of various experimental studies, that multiple-choice can be a viable auto-assessment method to improving student learning, especially for m-learning (de Marcos and Otón, 2010).

Luckin and Noss (2012) says that further consideration should be given to how technology can be used to enable the assessment of knowledge and skills not usually distinguished within current curricula. Clark and Killingsworth (2014) encourages a focus on how theoretically-driven decisions influence learning outcomes: for the broad diversity of learners, within and beyond the classroom. Combining these two, introducing e-assessment of entrepreneurship in a developing country context is a contribution to existing research.

2.3 Design for Learning

The following sections deals with how to design for effective learning (Dirksen, 2012). Designing for the mind is sometimes referred to as cognitive psychology, how our brain works in regard to mental processes such as retention and transfer of knowledge. There are a number of strategies and techniques that can be used when designing *learning* for the mind specifically: to decide educational objectives, how to reach these objectives (building skills and learning from assessment), and how to retain information (learning by reflection).

2.3.1 Learning Entrepreneurship: Mapping Educational Objectives with Bloom's Revised Taxonomy

What to teach should be determined by the learning objectives of the activity. Learning activities often involve both lower order and higher order thinking skills as well as a mix of concrete and abstract knowledge. All of this needs to be designed for depending on activity. Bloom's Revised Taxonomy (Krathwohl, 2002) separates between lower or higher cognitive complexity, and concrete (factual or conceptual) or abstract knowledge (procedural or metacognitive) (Cheong and Cheong, 2012). The taxonomy thus provides a framework for determining and clarifying learning objectives. Bloom's Revised Taxonomy is useful to map the whole range of learning objectives for entrepreneurship and as an entrepreneurship coach. See figure 2.1 from Heer (2012). Each colored block is an example of a learning objective matching with the two dimensions. The figure also explains the different concepts. Depending on the objective, it fits differently into the Knowledge dimension and Cognitive Process dimension of Bloom's Revised Taxonomy.

To craft good multiple-choice questions could be an art, but to map the question to learning objective makes it into more of a science: when designing for learning, using Bloom's Revised Taxonomy you can analyse if the educational objectives are met.

2.3.2 Building Skills: by Spaced Practice, Deliberate Practice and Perceptual Exposure

Spaced practice deals with spreading out learning, with the purpose of improving retention. Clark and Killingsworth (2014) concludes that spaced learning versus massed learning (no rest between sessions) did have a memory benefit in their study, and found no evidence of consistent correlation between total duration and effects on learning outcomes in their study. Taking spaced learning into consideration, could mean making the user apparent on the person's meta-cognitive ability (your personal insight of what you'll remember and when you are likely to forget).

To design for optimal learning outcomes of skills (e.g. entrepreneurial activity or coaching skills), deliberate practice has been proven effective. It has been tested before for mobile learning environments (Yengin and Uzanboylu, 2012).

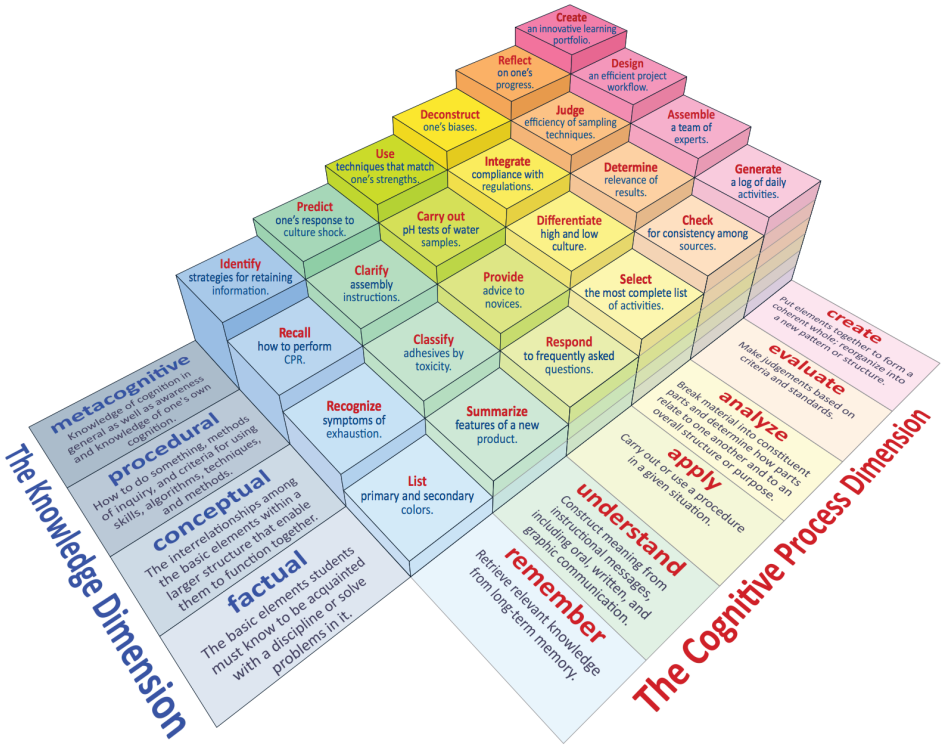

Figure 2.1: *Bloom's Revised Taxonomy visualised with examples of different learning objectives. The entrepreneurship topic question "What is financial literacy?" would be conceptual and remember.*

Building skills effectively, Sierra (2015) suggests helping users practice right, by designing practice exercises that will take a fine-grained task from unreliable to 95% reliability, within one to three 45-90-minute sessions. Sierra (2015) suggests skills to be divided into three buckets: can't do (but need to do), can do with effort, and mastered (reliable/automatic). The goal then is to move skills from can't do into mastered, in the best way possible. See figure 2.2 from Sierra (2015). If you can't get the user to 95% reliability within this time, stop trying; you need to redesign the sub-skill (Sierra, 2015).

Desirable difficulties applies here, meaning that during deliberate practice, it may feel as if learning gets more and more difficult, but in the long term the user is actually learning more. As a result, fewer people does true deliberate practice, but they do not get the same reward in return. Sierra (2015) suggests motivation to overcome desirable difficulties, see section 2.4.

By deliberate practice, you can practice better. The second attribute of becoming an expert is being exposed to high quality, high quantity examples of expertise (Sierra, 2015). It shows that whenever a skill relies on intuition, we could try exposing the user a well-designed trail and error test. In the case of multiple-

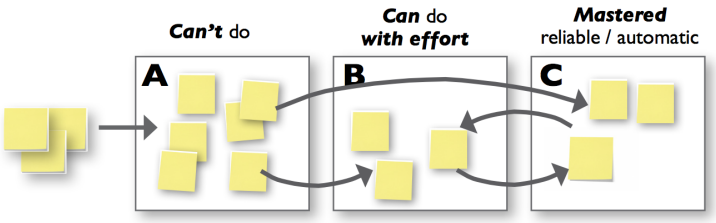

Figure 2.2: *Moving skills from A (Can't do) to B (Can do with effort) into C (Mastered) can move different ways, depending on how effective the learning is. Deliberate practices focus on A-B-C, while perceptual expose enables A to C. Reflection allows knowledge to go backwards, to get better at the skill than previously possible. An example might be to teach "Financial literacy". Concepts and factual knowledge (like what income and profit is) might need to move A-B-C, whereas entrepreneurship skills (like taking financial decisions) can move A-C if it becomes intuitive for the user, e.g. via having been exposed to a lot of trial-and-error examples in the app.*

choice questions, this could be done by exposing users to very high-quality samples during a very limited time. Perceptual knowledge includes teaching what we think of as expert intuition (like being a good entrepreneurship coach).

Sierra shows how researchers have repeatedly, by well-designed tests, been able to quickly build expertise by trial-and-error feedback. A novice would hazard a guess and an expert would say yes or no. Eventually the novices became, like their mentors, masters of the expertise that could otherwise would have been intangible for long.

2.3.3 Learning from Assessment

Regarding assessment, knowing what learners know and don't know is crucial to effective learning, Luckin and Noss (2012) says. Assessment can partly help to design for flow, as assessment can help perfectly match the learner's ability with the given challenge (Brühlmann, 2013), which is effective for intrinsic motivation (see section 2.4). Moreover, it also has cognitive benefits. It can help to offer appropriate feedback, increase learners' awareness of their learning needs, and give accurate assessment and analysis, and allows learning to be tailored.

By recognizing differences between students, in their ability to understand what they know and how they can progress, it is possible to ensure that everyone achieves their full potential. Effective assessment by a teacher or agent includes individual feedback (task-oriented and informal) and appropriate feed-forward advice. Sitzmann and Kanar (2008) has studied how questions used to prompt self-monitoring and self-evaluation benefit learning, showing gradual, positive effect on learning. Regarding multiple-choice tests, Nicol (2007) gives seven principles of good feedback practice, see figure 2.3.

Good feedback practice:
(1) helps clarify what good performance is (goals, criteria, standards);
(2) facilitates the development of self-assessment and reflection in learning;
(3) delivers high-quality information to students about their learning;
(4) encourages teacher and peer dialogue around learning;
(5) encourages positive motivational beliefs and self-esteem;
(6) provides opportunities to close the gap between current and desired performance;
(7) provides information to teachers that can be used to help shape teaching.

Figure 2.3: *Seven principles of good feedback practice Nicol (2007).*

Moreover, research on fixed mindset (I can't do X) versus growth mindset (I can't do X *yet*) talks about how mindset guides behaviour. In a mathematics game students were rewarded by the mentality of "Not yet" and effort versus getting a grade on existing knowledge (Dweck, 2014). Regarding learning, those exposed to a growth mindset mentality, previously having a fixed mindset, got superior results, especially those students previously having difficulties with learning. Regarding motivation, a study showed that high achievers played to the end regardless, but in the growth mindset version those still played to the end, but so many more lower and medium achievers also stayed until the end (Dweck, 2014).

2.3.4 Learning by Thinking: Reflection & Retrieval Practice

Stefano and Staats (2015) suggests that in some cases, learning from reflection is more effective than learning by doing. During the act of reflection, the student develops necessary skills and self-awareness to refine their own learning activities. His results suggest that reflection as an activity that can be more effective than additional learning. This surely applies to the teacher as well says Luckin and Noss (2012). Stefano and Staats (2015) found that individuals who are given time to reflect on a task, outperforms students who are given the same amount of time to practice with the same task. But, similar to deliberate practice, it is a desirable difficulty: individuals in the test themselves, had a tendency to believe that allocating time to practice on the task rather than reflecting on it would benefit them.

When it comes to study technique, Bjork (2016) as well shows that retrieval from memory is more effective than people who repeat reading the same thing to remember: the more effective students, retrieves from memory. One way to use memory retrieval as a study technique, is to ask "What was in that article?", before checking the answer in the article (the flashcard principle). It is an example of memory retrieval that is extremely effective for learning, their research shows. There is a danger with multiple-choice questions, that the student is given no time to reflect on the question and their prior knowledge, before evaluating the alternatives (Nicol, 2007). In the case of learning to teach entrepreneurship, where the information should be easily accessible during the teaching situation, selecting an alternative from multiple-choice does not simulate the teaching environment. Thus, memory via retrieval is essential in the multiple-choice design to have an effect in a real-world teaching scenario.

2.4 Design for Motivation

Research in motivation is commonly divided into three areas:

- Self-determination - the student is driven by a genuine wish that comes from themselves
- Achievement - the student is driven by extrinsic motivations, to be rewarded
- Expectation value - the student is driven by extrinsic motivations, but it is not coming from being rewarded

Koballa and Glynn (2007) and Abell and Lederman (2007) provide an overview of theories developed for these three areas. Further, Fulmer and Frijters (2009) provides a review of methods to collect data which can be used to study motivation. Deci and Ryan (1985) and Deci and Ryan (2000) studies self-determination theory, Elliot and Covington (2001) studies achievement theory, and Eccles and Wigfield (2002) and Wigfield and Eccles (2000) studies expectancy-value theory.

In terms of product development, Sierra (2015) argues that if you have designed for the user's compelling context the user is already motivated (by self-determination). Then, the motivation of the user is to become better (achieve). Sierra (2015) suggests the focus to be how to help users progress (see "Progress and payoffs", achieve), and what pulls them off (see "Cognitive load theory" below). See figure 2.4, for a summary of her focus on designing for building expertise.

In terms of using the product in a training environment, expectation value becomes relevant. Research within training transfer (Brinkerhoff, 2013) shows that before and after is as important as the training itself. To design for this, the leader should be involved with the participants before the training, and communicate expectations. The student should be expected to implement the training in everyday life (Brinkerhoff, 2013). Both of these aspects holds true for the YoungDrive program, and how a teacher communicates the use of an app needs to be taken into consideration.

2.4.1 Cognitive Load Theory

Sierra (2015) argues that working on what stops people matters more than working on what entices them. Thus, a focus needs to be identifying and removing stumbling blocks. She then describes how humans have scarce cognitive resources, and how to design for these. Cognitive load theory (CLT) research is divided into three types of load: intrinsic load (stumbling blocks), extrinsic load, and germane load (Sweller and Kalyuga, 2011). Below, to design for these are described.

Intrinsic load needs to be dealt with if the effort to perform a task is too high. Sierra (2015) describes two strategies. She first says that according to deliberate practice, if you can not get to 95% reliability within three 45-90 minute sessions, split skills that can be done with effort into sub-skills. The purpose is to reduce

Figure 2.4: *Sierra (2015) shares how to design for "making users awesome", meaning that a strategy for building an appreciated product is to help the user become good in something that appeals to her, i.e. to design for the compelling context. She then mentions barriers to not learning, like motivational aspects like reducing cognitive load and helping the user to progress.*

time spent practising being mediocre. Extraneous load, is about the way information is presented to a learner, and should be handled via designing to support cognitive resources, Sierra says (Sierra, 2015). Germane load, is the work put into creating a permanent store of knowledge. To make the knowledge permanent, make sure the learner believes the information is essential, Sierra (2015) says: either by designing for the compelling context, or designing for just-in-time learning versus just-in-case.

Scaffolding is a technique to step by step remove the support wheels for the user, e.g. present information in different ways. Clark and Killingsworth (2014) shows that in their research, each category of scaffolding demonstrated significant effects on learning. Another way to reduce cognitive leaks is don't make users memorise unnecessary things: make the thing you want the user to do, the most likely thing to do (accordances). Everything that takes willpower, reduces cognitive leaks.

The rule of thumb in applying cognitive load theory is that you want to decrease extraneous load, bad load, and increase germane load, good load. For example, in the case of a quiz app, the visual presentation of the questions can reduce extraneous load by removing unnecessary information. Germane load can be increased by giving scaffolding support at the end of each section, for example helping users to remember, by showing them their answers.

2.4.2 Progress and Payoffs

Sierra (2015) argues that to pull users forward, to stay motivated, progress and payoffs are essential. Both of these, are investigated in terms of motivational psychology.

The feeling of progress can be emphasised by a path with guidelines to help the user know where they are at each step, e.g. for a training. To create a path, she encourages the designer to make a list of key skills ordered from beginner to expert. Then, these are sliced into groups of ranking or levels.

This way, it is possible to design a “belt” path for your context. The first level, should feel like a superpower for the user. The best payoff, is a intrinsically rewarding experiences, according to Sierra (2015). For an entrepreneur, gaining the skills of selling (the progress) can be as rewarding as having gained the money for it (the payoff).

For motivation, the earlier, lower levels should be achievable in far less time and effort than the later, advanced levels. One practice is to try to have each new level take roughly double the time and effort of the previous level. This highly relates to flow.

Caring for the compelling context, why the user wants to learn the skill, are helpful strategies. A sometimes critiqued way of progression is to give the user high pay-off tips, but if done in a fair way, it is a good way for both learning and motivation.

This kind of path map is superb to simple gamification, says Sierra (2015). In an app for building entrepreneurship coach skills, the act of becoming certified (getting 100% correct), might not be as rewarding as the progress of getting there. Therefore, gamification of the sort "rewarding effort" might be more beneficial with "rewarding result". Suitable gamification could also mean unlocking new possibilities of adding value to the app (for example adding questions to the quiz), versus getting a badge or a star.

The shown benefits of designing for intrinsic motivation is in-line with self-determination theory (Deci and Ryan, 1985, 2000). Pink (2011) concludes that the surprising truth about what motivates us is that drive is fostered by autonomy, mastery and purpose. Meanwhile, Clark and Killingsworth (2014) says that simple gamification as well as more sophisticated game mechanics can prove effective. However, he adds that it should be investigated if "simple gamification" (e.g. contingent point and badges connected to learning activities) more frequently focus on lower-order learning outcomes, compared to studies with more sophisticated game mechanics. For the case of entrepreneurship, the goal is on higher-order learning outcomes, meaning simple gamification is not enough to motivate users. Thus, if you do not design for the compelling context, entrepreneurship coaches may well prefer other learning methods instead of using your gamified app.

2.5 Design Thinking in Multi-Disciplinary Projects

Interaction design talks about the creation of digital artefacts specifically (Löwgren and Stolterman, 2007). When it comes to the design process, it is influenced by related areas such as human-computer science, and more recently human-centered design. However, various disciplines suggest different design processes. For example, agile development suggest how do develop software efficiently. When-

ever a project is multi-disciplinary, various design processes may need to be combined. Whenever this happens, design thinking (how to think about design) becomes a skill essential to thoughtfully design the process.

Löwgren and Stolterman (2007) writes about design thinking and useful techniques in general, from his interaction design perspective. Service design thinking connects various fields of activity (Stickdorn and Schneider, 2010), and it's methodology relies on being close to the users. While interaction design talks about the creation of digital artefacts specifically, service design talks about the creation of services. As some digital artefacts are used within a service, or can be thought of as both a product and service simultaneously, the combination of the two can be very useful. Service design could help the designer be aware of how such a artefact would need to interplay with its physical environment.

Each discipline holds efficient methods and tools, that can be modified to suit the specific situation even better. From the field of graphic design, mental models describe the perceptions of the user. From interaction design, desirability, utility, usability and pleasurability can be useful principles to evaluate a product (Clatworthy, 2010). While none of these are a mandatory part of service design, these have been useful in service design projects previously (Stickdorn and Schneider, 2010). In difficult situations, combining different disciplines places demands on the designer. This is where design thinking becomes relevant. Below, relevant methods and tools are briefly described, and what it means to be a good designer.

2.5.1 A Good Designer

The result of a method can not be better than the people engaging in carrying out the process (Löwgren and Stolterman, 2007). With its user-centered focus (Stickdorn and Schneider, 2010), service design can be said to equip the designer with tools both for reasoning and design ethnography. But it also suits to get to know and design for the learning situation. In learning, the end goal is that the student raises their level of knowledge and expertise, and the design needs to be adapted for this specifically. Central to design for learning is to dig deep into the topic being communicated. In this case, understanding entrepreneurship, understanding exactly what is being taught (the training), and adapting the design after this.

A good designer can deal with the complexities of design: a satisfactory (and surprising) solution or design can be achieved while working in a highly restricted situation (Löwgren and Stolterman, 2007). This can be done e.g. by inventing new design techniques. One such example that would suit designing an app for entrepreneurship training in a development country, would be a *field hackathon*. A field hackathon would thus allow that during the training, the topic *and* the users are observed and understood. Then, the app can be tested (in this can a quiz assessment of the trained material). Then, users can be invited to give feedback, suggestions of improvements, and ideas. For the next day, an improved version of the app is tested, and then the process is repeated.

More examples of how a service design process can be invented to deal with digital artefacts, can be described in the chapter 2.7. However, to do such field tests (like a field hackathon), requires building trust and having an enabling en-

vironment, which is where relationships and roles become crucial.

2.5.2 How to Deal with Relationships and Roles

While a researcher is interested in reality, a designer is interested in what reality could become (Löwgren and Stolterman, 2007). Being thoughtful means conceptual clarity from the designer, caring for the vision, and being equipped with appropriate tools of reasoning. These are all good characteristics for a successful project. According to Löwgren and Stolterman (2007), "real" design is about finding ways to design a project within the existing preconditions and limitations. Being innovative and communicating well with the stakeholders becomes crucial.

There are three roles an interaction designer in particular can take: the computer expert, the socio-technical expert, and the political agent. The trend is increasingly towards socio-technical experts (Löwgren and Stolterman, 2007), the middle ground, as human understanding and collaboration is so important. This seems to be a perfect fit with service design, where interaction design is both technical skills and design, and service design can be both design and ethnography. Even more importantly, service design suggests making the whole process co-creative, involving all stakeholders (Stickdorn and Schneider, 2010).

2.5.3 Thinking of a Product as a Service

Service design thinking is described as a process of designing, rather than to its outcome. A service's intent is to meet customer needs. If it does, it will be used frequently, and recommended (Stickdorn and Schneider, 2010). As this is often not the case, service design can be applicable to fields including social design, product design, graphic design and interaction design. The result can be a product service hybrid. When designed and considered well, service design shapes the value proposition and desirability of the product for the better.

2.5.4 Starting the Project

Löwgren and Stolterman (2007) writes about the beginning of a project: This is where the designer gets involved in design work, establishes a preliminary understanding of the situation, navigates through available information, and initiates all necessary relationships with clients, users, decision makers, and so forth. Based on all this, she creates a design proposal (Löwgren and Stolterman, 2007).

2.6 Service Design Methodology

Below, brief descriptions of five principles of service design are described according to Stickdorn and Schneider (2010), together with how the work is divided into iterations, and examples of tools that can be applied.

2.6.1 Principles

Stickdorn and Schneider (2010) describes five principles that constitute service design thinking, and how to follow these.

He describes how to follow these principles, by making the process user-centered (e.g. via *design ethnography*), co-creative (involve all stakeholders) and holistic (keep the big picture). Sequencing (visualize the service, and make iterations) and evidencing (make the service tangible) are the two last important principles.

2.6.2 Sequencing

Sequencing the process means splitting the design process into iterations, which consists of a number of steps, which are repeated for each iteration. This is a common denominator with the agile methodology SCRUM, which is often applied in software development.

While service design literature and practice refer to various frameworks, regardless of number of steps, every service design project includes: exploration, creation, reflection and implementation (Stickdorn and Schneider, 2010). Widmark and Nissar (2016) suggests a model where one iteration consists of insights, ideation, trigger material, and interactions. See figure 2.5.

1. Interactions, where you are listening, the *Explorative phase*.
2. Insights, which is where you use the Interactions in order to try to understand, the *Understanding phase*.
3. Ideation, where you find possible ideas and when creation of a new version of the app is done, the *Design phase*.
4. Trigger material, where material is developed to test the outcome of our evaluation in the next round, the *Trigger development*.

The iterations should come closer and closer to a desired outcome. It is not always obvious what this outcome is. For each iteration, the process takes the project closer, from Why? to What? to How?, often with overlaps (Widmark and Nissar, 2016). See figure 2.6.

2.6.3 Service Design Tools

There are a number of popular service design tools that follows the five principles, e.g. how to make it user-centered. One is Customer Journey Map, in which an activity (like hosting a youth session) is broken into Before, During and After. Another method is Personas, which *exemplifies* thought users of the app into people with names, having realistic character traits and opinions. The persona's needs can then be thought of when designing. An alternative to Personas are Need groups, where thought users are broken down by their different needs. Instead of designing for a specific person, you design for a person with a specific need. The advantage of Need groups, are that it accepts the view that the same

Figure 2.5: *In the model by Nissar (2016) an iteration consists of Interactions, Insights, Ideation and Trigger material.*

person (Persona) might have different needs, depending on the situation. The 5 Why's is a simple method used to dig deep into understanding the interviewee. Variants of the question "Why?" is repeated five times as a rule of thumb, to understand underlying motives. This method is called "Why-why-why" within interaction design (Löwgren and Stolterman, 2007)).

Tools to create and reflect can be done via certain work methodology. When you structure and inspire brainstorms, you can ask "What if...?" and do Co-Creation, meaning doing ideation together with stakeholders or users. To create, agile development can be used, which is often suitable for software engineering. The manifesto for agile development is (Alliance, 2001):

- Individuals and interactions over processes and tools
- Working software over comprehensive documentation
- Customer collaboration over contract negotiation
- Responding to change over following a plan

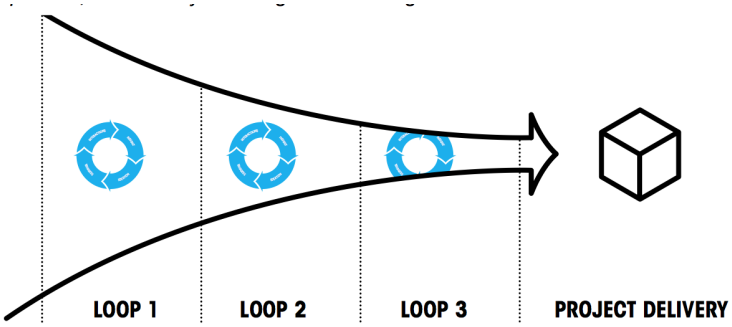

Figure 2.6: *The iteration process consists of a number of iterations with different focus, starting with broad strokes, and narrowing down into a concrete product. Between iterations, there is an overlap in "Why?" and "How?", "How?" and "What?", which signals that there is a learning process which means conclusions may need to be quickly questioned as new insights emerge. This is especially important in projects where you work with an unfamiliar target group and there are several uncertainties and constraints.*

An example of an agile methodology is *SCRUM*, where a project is divided into several iterations similar to a service design approach of sequencing, but also introducing concepts like retrospectives (reflecting on one's work) and sprint demo (demonstrating the results of the iteration to stakeholders) (Kniberg, 2005).

There are also some service design best practices: interviews are often done via open questions (encouraging stories) and dialogue can be facilitated with a questionnaire guide. In workshops, post-its are often used, and followed up with specific questions. Service design methodology encourages taking pictures, filming and recording audio, benefiting the analysis done afterwards (Widmark and Nissar, 2016).

2.7 Digital Service Design

The method combines the benefits of Service Design, Agile Methodologies (namely SCRUM) and Interaction Design. Its purpose is to contribute a holistic approach to the digital design solution for a specific target group. The methodology was co-created by the current author and Expedition Mondial for the master thesis (Nissar, 2016).

2.7.1 The 4 stages of a "Service Sprint"

In Digital Service Design, an *iteration* is called a *service sprint*. Each iteration includes four steps: insights, ideation, trigger material and interactions. Each step borrows a number of best practices from agile development or interaction design. There are also new methods, like how a *field hackathon* includes *mini service sprints* each day. Below, the four steps are presented.

2.7.2 Step 1 - Insights: Analysis, Retrospective & Stakeholder feedback

Insights consists of *analysis* (service design), but also a *retrospective* (SCRUM) and *stakeholder meeting* (service design). In the analysis, the app is evaluated (in terms of interaction design - pleasurability, usability, utility and desirability (Clatworthy, 2010)), and quantitative data is processed (often by clustering data points) and compared with qualitative data (quiz results and questionnaires). This produces an analysis overview of the result. In the *retrospective*, the design process is evaluated ("start doing, stop doing, continue doing"), and changes to the design process are suggested for the following iteration.

Both the result analysis and the design process analysis is then presented during two stakeholder meetings (service design), structured as *sprint demo*s" (SCRUM), with the purpose of getting feedback. The first "Expert meeting" informs the next iteration's design process, while the second "Partner meeting" informs the next iteration's delivery. From the new insights, a *product backlog* (SCRUM) is converted from needs and ideas into *stories* balancing 1) user needs and 2) stakeholder needs.

2.7.3 Step 2 - Ideation: Planning Interactions and Delivery

Ideation consists of doing *sprint planning* (SCRUM) for the trigger material (a *lo-fi* or/and a *hi-fi prototype*) and the interactions (where tests and workshops and field visits happen).

- Trigger material
    1. Ideas are formulated which would satisfy the user needs. This is often an iterative process, which happens in dialogue with chosen experts and entrepreneurs in technology, design and education.
    2. To plan the implementation of the ideas, every technical task is laid out, measured in time and prioritized. The least prioritized tasks can thus be cut or moved to the next iteration, in case it is necessary.
- Interactions planning
    1. If the technical planning has been realistic, it is time to determine what this iteration's interactions should look like. How will this be tested?
    2. The interactions activities are chosen (what, how, when), so that these are communicated to the local partner, who may schedule the days that will be visited, and solves the needs to the best of their ability.

2.7.4 Step 3 - Trigger Material

Trigger material is about preparing the interactions (field visits, interviews, app tests, workshops) and creating the low-fidelity (pen and paper) and high-fidelity prototype (developed app) to be tested with the users. To track the progress and

plan effectively, each day starts by a daily stand-up, where today's targets are set, ending by reflecting if the targets were met. If they were not, either the design process needs to change, or something needs to be cut short.

2.7.5 Step 4 - Interactions: with "Service Mini-Sprints"

Interactions always consist of a sprint demo with the users with the low-fidelity or high-fidelity prototype. During the development process, these are *formative tests*, while for final app evaluation, this is a *summative test*. **Group tests** are facilitated as workshops. Often, a scenario is presented, devices are given, results are submitted, followed by an open discussion. **Individual tests** are facilitated in the field (using the before, during, after technique). I observe how the coach does the job today, tests and observes if the app fits into the process, followed by an interview.

These tests always inform what steps to be taken next, both in terms of app development and interactions. Instead of waiting for the next iteration to do these changes, often a so-called Service Mini-Sprint is done.

In a **service mini-sprint**, the insights gathered during the day allows for last-minute adjustments of coming pre-planned workshops (*co-define*, *co-create* or *co-refine*) or field visits (change of interview questions), that can sometimes happen the same day. To take advantage of the precious time with the coaches, at the end of the day, app improvements are made and tomorrow's design process revisited. This means, that already the next day, an improved version of the app can be tested. Similarly, improvements to a workshop format can be improved. These mini-sprints allows for very fast iterations, which can sometimes accelerate the outcome of the visit.

2.8 Hybrid App Development

The history of app and web development is rich and increasingly intertwined. First, websites were developed for desktop only, and when smartphones became popular, they were made responsive. With today's possibilities of native mobile development or developing a native app using web technologies, there are numerous viable alternatives available if an app should function on several devices, depending on budget and preferences.

One of the main argument for developing an app in web technologies, is that the whole application, including the server, can be written in one programming language, JavaScript (full-stack). Tools such as Cordova (Apache, 2011) can then compile JavaScript applications into native apps. Thus, they can appear on Apple iOS and Android Play Store, as well as on the web, or installable offline on a smartphone from the computer.

JavaScript is developing rapidly as a language, as well as its ecosystem of frameworks and tools. Frameworks have emerged and matured, like Meteor (Inc., 2016), which makes building full-stack applications in JavaScript reliable and fast. Previously, web hosting has been troublesome for JavaScript server applications. Today, tools such as Heroku (Salesforce, 2007) have introduced free and

paid hosting for such applications, with smart bindings to code platforms such as GitHub (GitHub, 2008), which makes collaboration and version handling easy.

3

Methods and Implementation

To understand the users of the app and the design possibilities, the setting and research context is described, together with a description of the participants. Subsequently, application implementation is described, followed by presenting the study design and data collection. A data analysis framework presents methods used to analyse quantitative and qualitative data.

3.1 YoungDrive, Terminology and Limitations

In this section, the organizational structure of YoungDrive and the master thesis is described, also explaining why there is an opportunity to innovate in regards to mobile technology in Uganda.

3.1.1 Entrepreneurship in Uganda

A couple of prominent facts helps to portrait entrepreneurship activity in Uganda:

- Uganda can be ranked the world's most entrepreneurial country (with 28% of the population considered entrepreneurs (Lönn, 2015a))
- Uganda is ranked the youngest population in the world (78% of all Ugandans are below 30 (Balyesiima and Kaggwa, 2012))
- Uganda has a very high unemployment rate (62 % of people between 18–30 are unemployed (AAU, DRT, UNNGOF, 2012))

With a high unemployment rate and little or none social security, starting a business is for many young entrepreneurs simply a tool for survival. But tough conditions can also lead to creativity, and there are as well many innovative entrepreneurs with great ideas and the aim to create positive social impact (Nissar,

2016). As Mitchell and Smith (2007) says about entrepreneurship, the motivation of entrepreneurship does not need to be solely wealth accumulation anymore. The activity of entrepreneurship contributes to society, in a way that is not captured by the commercial entrepreneurship literature. No matter the reason of starting a business, Uganda's many entrepreneurs are contributing to the national society by boosting the economy and creating new jobs.

3.1.2 YoungDrive

YoungDrive (2016) is based on a Swedish concept, and had previously had a pilot in Botswana, when tasked with running the entrepreneurship module of A working future for Plan International. The organization fosters and educates young entrepreneurs in developing countries. They train the coaches, provide training material, and support the coaches via direction and direct support through co-project leaders and Youth Mentors.

YoungDrive moves an entrepreneur to location, becoming country manager and "teacher". The teacher educates project leaders during four days, followed by educating coaches, which then roll out the training to the youth groups during 10 sessions, 1 session per week in average. The Community Based Trainers (CBTs) also rolls out other trainings, often simultaneously.

For the future of YoungDrive, they want to make the CBTs even better, and collect and take use of data (monitoring and evaluation). Another motivation is scaling and monetization, as Plan International wants to increase the project to more countries, with an increased digital focus, and YoungDrive wants to be independent of project funding (i.e. a social enterprise).

Given the above, this was a great time to introduce digital enablers for YoungDrive, where there previously had been no technology-focus, especially towards CBTs and Youth Mentors. The master thesis is the first project which focuses on digital enablers for YoungDrive.

3.1.3 Roles within YoungDrive

The *country manager* trains the project leaders. It is also the main person responsible for partnerships and the quality of the YoungDrive program in the respective country. In Uganda, the country manager is Iliana Björling. She is located in the Uganda capital, Kampala, which is a strategic location because it is the same city in which the national office of the main partner, Plan International, is located. In Zambia, the country manager is Josefina Lönn, who previously was the project leader in Kampala, and has held all the trainings up to this point. Now, she leads the operations and has trained the coaches in Zambia, in the new role of country manager and project leader.

The *project leaders* train the coaches and oversee the coaches, manages the coach training, and also collaborates with local stakeholders for quality assurance and to oversee daily operations.

The *coaches* train the youth. In Zambia, a coach only has responsibility for training youth in the YoungDrive program. In Uganda, this is called a *Youth*

Mentor (YM), in contrast to being a *Community Based Trainer (CBT)*, which also trains the youth in other programs and leads the youth saving groups. Most of the CBT's in Uganda holds sessions together with a Youth Mentor, or divides work between them, instead of being alone. The coaches are often volunteers, receiving a small scholarship from the partner organization. They are often business owners themselves. The coaches could be described as social entrepreneurs (Mitchell and Smith, 2007). Many of the YoungDrive coaches (and youth) are driven by that their business can have an impact on their community, *as well* as take them out of unemployment or increase their current livelihood.

The *youth* are the ones receiving the training from the CBTs and the YMs, being encouraged to start their own businesses.

3.1.4 Mobile Technology in Uganda's Rural Areas

One of the reasons why mobile services are growing rapidly in Uganda is that the country has invested heavily in communication networks, even connecting remote rural villages with fibre optic cables and thereby connecting them to a world of information.

As much as 65% of the adults in Uganda owns a cell phone, which has allowed many areas in the country to skip the landline stage of development and jump right to the digital age. For those who hasn't electricity at home, there are available charging booths for mobiles all over the country.

The wide use of mobile phones in Uganda and other developing countries has lead the way for the development of several innovative mobile services and in many cases the mobile services are way ahead of us (Nissar, 2016). In Sweden mobile banking services that allows us to transfer money through our mobile phones were made popular with Swish, introduced in 2012. In the neighbouring country Kenya, people have had similar services for the last 10 years, and mobile money is since long also common in Uganda.

A prominent example of an app that has previously been developed with the target group in mind is Ledger Link (Foundation, 2014). This mobile banking service empowers, developed in partnership with a bank, allows saving groups in rural areas such as Tororo to save money remotely. It is developed with human-centered design methods, and has won several awards.

3.2 Collaborators for the Master Thesis

Collaborators in the project are the current author, supervisors, stakeholders and experts. Below, the responsibilities of these are more clearly laid out.

The Current Author

It is needed to take on several roles in the project by the current author: most notably that of a project leader, designer and developer. It is needed to balance stakeholders' different opinions and requirements, and caring for the vision in order for the project to be successful (see section 2.5.1 A Good Designer).

There are two groups, with the current author included in both of them, which gather at the end of each sprint for a check-up meeting. The Expert group consisted of Expedition Mondial and LiU Innovation. Expedition Mondial could help with the design process, and LiU Innovation could offer input on social innovation. The meetings mostly lasted for one hour. The Partner group consisted Iliana Björling from YoungDrive, and Lena Tibell and Konrad Schönborn from Linköping University. In Partner meetings, the insights from each iteration was presented and discussed. Then possible decisions were laid out, followed by discussing the alternatives. Outside of these groups, these people can also give advice in certain situations. For specific areas, there are also some experts which have been beneficial during the projects. Below, the whole team is explained:

Supervisors

The supervisors are from YoungDrive and Linköping University. The YoungDrive team consists of Iliana Björling, founder of YoungDrive, and Josefina Lönn, country manager in Zambia. They are both helpful in giving knowledge on the entrepreneurship education program, and giving support. The Linköping University team consists of Lena Tibell, Professor, and Konrad Schönborn, Doctor, within the Department of Visual Learning and Communication.

Stakeholders

The stakeholders are considered YoungDrive and Plan International. **YoungDrive** is the client of the work, and their needs should be satisfied. This person is mainly represented by Iliana Björling, who is part of the YoungDrive Strategic Management Team. Using service design, the project leaders in Uganda and Zambia, are also considered stakeholders: Josefina Lönn in Zambia, and the two co-project leaders in Uganda. Finally, the most important stakeholder of all according to service design, is the actual users: the coaches. They should be the main consideration of the work.

Plan International is the organization allowing for all the interactions with the end users in Uganda. A similar organization is operational in Zambia. They are the ones that are providing facilities, organizes transport, etcetera. They in turn, have the organization Community Vision, which organizes the coaches. If Plan International or Community Vision does not appreciate of the project and the collaboration, then the interactions with the coaches will not be possible.

Advisors

Since the development country context is new to the current author, there are also specific experts advised in the project. For design process, Susanna Nissar and Erik Widmark from Expedition Mondial has supported with all of their knowledge within service design. Julien Tantege, Research Specialist at Grameen Foundation, has been kind to offer support before and during the work, sharing their insights from related work, and giving feedback during ideation. She has experience doing technical development for rural areas. For pedagogical development,

Henrik Marklund from edtech startup Knownly in Sweden has given support with regards to building skills within digital learning. For feedback for how the work relates to social innovation, Peter Gahnström at LiU Innovation has offered feedback.

3.3 Describing the YoungDrive Coaches and the Research Context

The biggest challenge with regards to time constraints and cultural differences is that it is difficult to understand the target group of the app. Therefore, the whole design and development process will take place in Uganda, with several interactions with the intended users. The work was carried out from Hive Colab, a co-working space and an innovation hub. The work is done mainly from Kampala, because that is where YoungDrive is situated, meaning that there is still a long distance to the coaches and youth in Tororo, which is located near the Kenyan border. Another challenge with being in Uganda compared to Sweden is that internet speed and access is worse, especially outside Kampala.

The interactions took place in either Uganda or Zambia, in the locations where training of the coaches and youth takes place. There were a number of resources made available to support the work, for example the YoungDrive manuals. Each youth is given a *Participant manual*, describing each week of the 10-week YoungDrive program. Coaches are also given a *Coach guide*, which describes how to carry out and teach each week's topic during the youth training. As most of the coaches did not have smartphones or tablets, four smartphones (3 Android, 1 iOS) and ten tablets (3 Android, 7 iOS) were brought from Sweden. All of these devices had a web browser and access to an app store. These were either donated, borrowed or bought devices. During the user tests, also using a laptop would be tested. The following section describes the Ugandian and Zambian coaches and their businesses.

3.3.1 Social Characteristics and Businesses in Uganda

According to statistics gathered by YoungDrive during 2015 evaluations (Lönn, 2015c), there are a number of considerations to make regarding the coaches in Uganda. This regards entrepreneurship experience, technical access, and language, see a summary in figure 3.1. Taken together the coaches' and project leaders' technical skills are currently low, and this needs to be in consideration when designing the app. Regarding language, English can be used in the coach app.

The coaches in Tororo are divided into three different regions. Based on region, income and experience, they run different kinds of businesses. [1] In Tororo, the coaches' businesses range from: pineapple, water melon, onion, chili, bakery, catering, corn, beans, fabric, plastic products, bird farm, milk, fish, ground

[1] In Uganda and Zambia, a small-scale business is typically not registered. Thus, the coaches' definition of a business can be more generous.

	# Roles	# Responden	% Respondents	Running business	Has cell phone	Internet on phone	Electricity at home	Solar energy at home	Can write on computer
Project Leaders (Zambia)	1	1	100%	100%	100%	100%	100%	0%	100%
Project Leaders (Uganda)	6	6	100%	100%	100%	17%	33%	50%	100%
Tororo Coaches	27	26	96%	100%	100%	12%	0%	12%	15%
Tororo Youth	6000	225	4%	73%	44%	5%	2%	13%	12%

Figure 3.1: *Table showing entrepreneurship experience, technical access, and language between Tororo coaches and the project leaders in Uganda and Zambia. All of the Tororo coaches run a business (with a majority running more than one). This means they do have practical experience of running a business outside of the YoungDrive coach training. While all have a cell phone, smartphones are very uncommon - only 3 uses Internet on the phone, every day or weekly (mostly for Facebook or email). Regarding power, none has power at home, 3/26 has solar, and only 4/26 can write on a computer. While about half of the asked Uganda youth can not understand (129/225), read (133/225) or write (132/225) English, most of the coaches in Uganda are proficient. These characteristics are similar for youth and project leaders as well.*

.

nuts, cabbage, tomato, hairdresser, sewer, shop and rice. For photograph of the environment, see figure 3.2.

Figure 3.2: *One of the co-project leaders showing the rural part of Tororo, where crops are growing close to where he lives.*

.

In Tororo, there are 2 Project Leaders. one project leader's business range from: bakery, corn, pig farm and plastic products. The other person's business range from: silver fish, beans, corn, and bird farm. In comparison, in Kamuli, there are 4 Project Leaders. Their businesses range from: selling office supply, motorcycle taxi, bird farm, pig farm, green pepper, corn, cabbage, tomato, aubergine, chipati ("bread"), chilli, and charging of cellphones.

Among the youth, the top 8 most popular businesses in Tororo, with 134 respondents, are corn, cassava ("potato"), saloon, fish, making of bricks, beans, brooms and rope. These range from 9 for corn (6.7%) to 5 for rope (3.7%).

3.3.2 Social Characteristics and Businesses in Zambia

During the visit in Zambia, the coaches had not yet formed their youth groups, and started their own teaching. Regarding characteristics, ages ranged from 21-39 years old (26.8 average). Other data available about the coaches were notes of the Zambian coach job interviews, which could be compared with during quiz analysis (Lönn, 2015b). Compared with Uganda, 9 out of 10 had business experience.

3.4 Study Design and Data Collection

As a computer expert with social skills needing to design and develop an app for an unfamiliar cultural and socio-economic context, it was needed to quickly become a good designer. The technical aspect of the project was but one. It was needed to learn how to develop hybrid apps in JavaScript that worked offline, and had an online back-end. However, those are merely the technical demands.

It was needed to quickly become a good designer, not mainly from a perspective of graphic design or interaction design, but *how* to explore, design, and implement what the user needs from the requirements "fun, user-friendly, and good for learning". There was also a need to evaluate the effectiveness of the implementations, to assess learning and the interaction design aspects of desirability, utility, usability and pleasurability. The approach used to learn design from these perspectives was to read extensive literature, consult a diverse set of experts, and be humble and curious in interactions with the end-users and stakeholders. In the following section, the creation and implementation for a suitable design process is described, together with the study design and data collection for each iteration of the project.

3.4.1 Creation of Design Process

As there was a unfamiliar target group - mostly young Ugandans with little or no experience of smartphones - service design thinking would benefit true understanding of cultural context and in-depth empathy for the end users. Tools and methodology in service design were chosen with the help of Expedition Mondial in Stockholm, who provided education and coaching.

The end result would be a digital artefact (an app), which is not common in service design. While this product could be thought of as a service, the tools and methodology would benefit to borrow from Agile methodology and Interaction design. Being a computer expert kind of designer (Löwgren and Stolterman, 2007) helps being adjusted to agile methodology and interaction design. However, aspiring to be a socio-technical expert, Expedition Mondial are consulted, who are experienced with service design, and are aspiring to be more of computer experts. This led to the joined development of a Digital Service Design method, co-created by the both them and the current author. The result is that the design and development phase in Uganda is an iterative process with the human in focus. The process is built on top of service design process and methodology, while

in-line with digital design practices.

3.4.2 Implementation of Design Process

There were four iterations, going through the stages of Insights, Ideation, Trigger Materials and finally Interactions with the coaches. For detailed time plans, see appendix C. The first iteration follows Service Design Methodology, with no app development, while the other three follows the new methodology, Digital Service Design, when starting to construct the app. In iteration 1, there is a very broad focus, to understand the limitations and possibilities of developing an app for the YoungDrive coaches. Iteration 2, 3 and 4 introduces and narrows down the project into a digital solution. In Tororo iteration 1 it was chosen to observe the youth sessions. In Zambia iteration 2 it was decided to observe the coach training. In iteration 3 in Tororo it was chosen to observe the coaches preparing the youth sessions. Figure 3.3 is made to assist the reader in the iterations' different foci. Figure **??** is made to assist the reader in the study design, which activity led to which data collection and data analysis for each iteration.

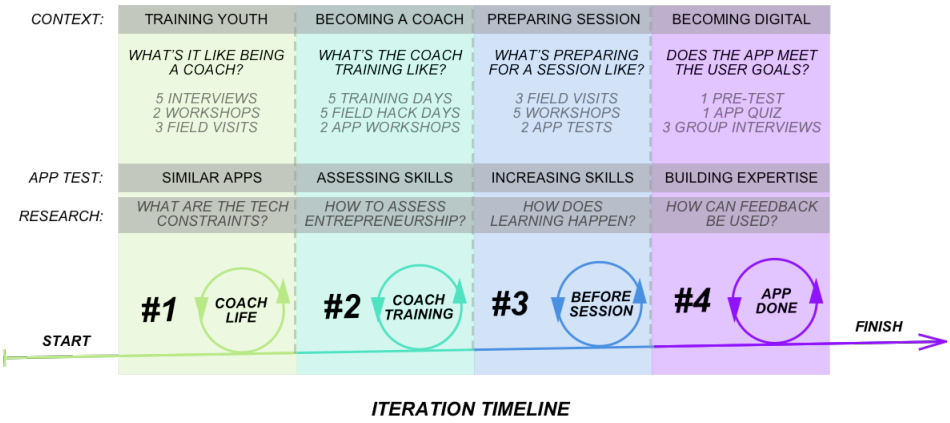

Figure 3.3: *Each of the four iterations had a unique context, app test focus, and research focus. The loops are meant to remind that an iteration consists of three steps before Interactions with the coaches: Insights, Ideation and Trigger material. For more information, see section 2.7 Digital Service Design.*

Expedition Mondial gave support in each iteration, helping with refinements of each iteration as learnings happened along the way, and they were able to educate the current author during the different stages with methodologies whenever necessary, see figure 3.5. The supervisors from Linköping University helped with how to construct the app tests for each iteration, and assisted with the Bloom analysis of the questions.

	ITERATION #1	ITERATION #2	ITERATION #3	ITERATION #4
ACTIVITIES:	Stakeholder interviews Coach interviews Workshop Smartphone test Field visits	Field visit App tests Workshops	Field visits App tests Workshops	Pre-test App quiz *- Test or reference group* *(- with or without manual)* Group interviews
DATA COLLECTION:	Notes from interviews, workshop, smartphone test and field visits	Test items Quiz responses Log data Notes from field visit and workshops	Quiz responses Log data Notes from field visits and workshops	Quiz responses from test and reference group Log data Notes from interviews
DATA ANALYSIS:	Analysis of notes*	Bloom analysis of test items Analysis of quiz responses Analysis of notes*	Analysis of quiz responses Analysis of notes*	Analysis of quiz responses *- Statistic analysis* *- Parallel coordinates* Mind-map analysis of evaluation interviews Analysis of notes*

* Personas/Need Groups, Customer Journey Map, Sprint Backlog and Sprint Demo

Figure 3.4: *Overview of what activities, data collection and data analysis was done for each iteration.*

3.4.3 Iteration 1: Uganda Coach Visit

Following the service design sequencing, the first iteration had a very broad scope and truly is a service design iteration: "From your perspective, what is it like being a coach?[2]". Starting the project had the goal of getting a preliminary understanding of all important aspects, and build relationships with all stakeholders.

All insights depended heavily on interviews with all the stakeholders (2 with Plan International, 3 with YoungDrive), and local experts (1 visit each at Grameen Foundation and Designers without Borders, 1 workshop with Mango Tree), since no Interactions with users had been made yet. Also, knowledge and connections were made with the Kampala technology scene as much as possible, which was benefited by working at the technology hub and co-working space Hive Colab.

Ideation was about creating a questionnaire guide for the interviews, a co-creation workshop using "Customer Journey Map", and identifying how the app test should be designed to test their existing knowledge (and be informed of the design preferences of the YoungDrive app).

Trigger material was the finished questionnaire guide (constructed with Expedition Mondial) a written plan for the co-creation workshop ("A day as a coach"), and a written plan for testing the quiz app Quizoid (Apps, 2016) and the language learning app Duolingo (Duolingo, 2016), and a schedule for the interactions.

The interactions were focused on design ethnology, getting to know and learn from people in a different culture, namely the coaches. The focus was on their needs, motivations, and environment. To understand these, four days were spent in Tororo, with one day of travel. There were four face-to-face interviews, one meeting with Plan, one meeting with the local partners, two workshops, one

[2] A coach meaning either a Community Based Trainer (carrying out all the trainings), or a YoungDrive coach, depending on who was asked the question.

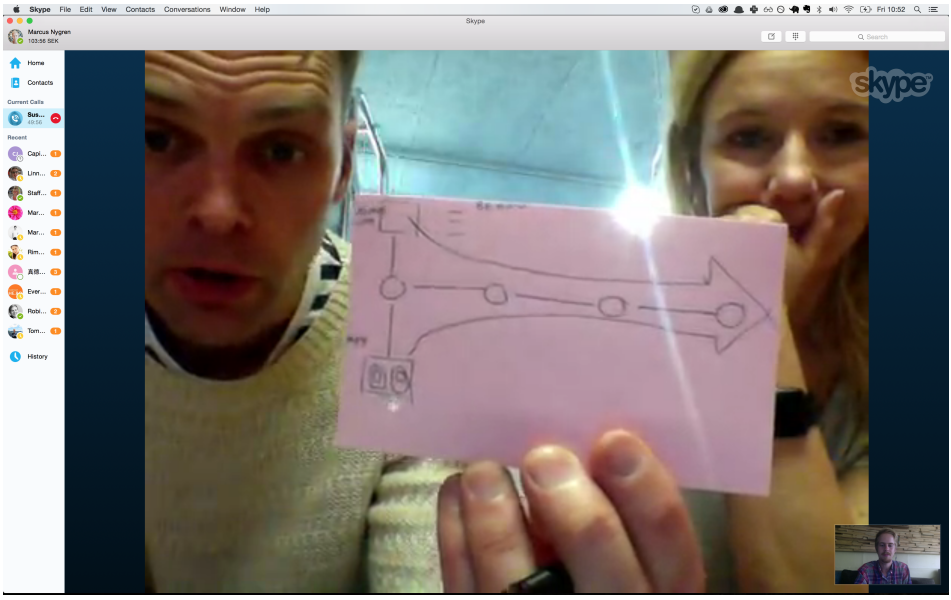

Figure 3.5: *In preparation for new situations, like meeting coaches and doing interviews, Nissar and Widmark gave support via Skype videos, sharing service design methodology tips or giving feedback on ideas.*

coach stay-over, and two youth session visits (one of the youth sessions are observable via figure 3.6.

Customer Journey Map: "A Day As a Coach"

The first workshop had the tree interviewed coaches as participants. See figure 3.7. A Customer Journey Map would be created, with the purpose to understand all activities involving “A day as a coach”. The structure of the timeline was: “Before”, “During”, and “After” a youth session. Too understand how these activities differentiated between different coaches, three Personas were co-created based on the previously discovered types of coaches: "the ideal coach" ("John"), "the realistic coach" ("Joan"), and "the challenged coach" ("Suzan").

3.4.4 Iteration 2: Zambia Coach Training

This time, the iteration has a more detailed scope, with a hypothesis on what needs the app should meet in the end, and create low-fidelity and high-fidelity trigger material to meet those needs. A co-creation workshop started the interactions, followed by repeated app tests at minimum one session per day, always followed by a feedback round, so the app and the tomorrow’s question set creation could be improved for the next day. At the end of the week, there was a co-refinement workshop of the current high-fidelity material, and also low-fidelity

Figure 3.6: *Photograph from the location where one of the youth sessions were held. Here, a YoungDrive coach in Tororo teaches youth in basic sales and marketing.*

Figure 3.7: *Two local project leaders and one coach mapping out the activities a coach does before, during and after hosting a youth session, in what is called a Customer Journey Map.*

material for the new version of the app.

Creation of Questions

Project leader Josefina in Zambia refined Iliana's first question sets, prepared for the visit to Zambia. Josefina created question sets mostly taking into account the structure and the order of the coach manuals, what it means being a coach within the topic, and lastly scenarios, but also thinking about assessing a variation of levels on Bloom's Revised Taxonomy. After the interactions, the questions created from were analysed using Bloom's Revised Taxonomy in partnership with the supervisors from Linköping University.

Trigger Material Used

A high-fidelity trigger material was done, a very basic quiz app, keeping it as simple as possible (see Application Implementation, Iteration 2). All of the devices

(tablets and smartphones) that were kept in Kampala were brought to Zambia. Josefina's questions were added to the app, and the app was installed on all of the devices. This process was repeated for all the days, Sunday-Friday.

Design Workshop #1 in Zambia

The coach training started with having a design workshop with all (10/10) of the coaches, see figure 3.8. The co-creation workshop was made to let the coaches create a low-fidelity prototype, and as a result identify important functionality in the minds of the coaches, before being biased by seeing the high-fidelity trigger material created in advance.

1. Since the experience with smartphones and apps were mostly non-existing, an introducing to these topics was made.
2. All were familiar with what Facebook was, so thus the Facebook app was shown as an example. Wanting to know what the app would look like if the coaches would have designed the app, they were trained how to design an app via drawing wireframes.
3. Using post-its, they started with during limited time drawing the start view from the Facebook app.
4. Then, they were asked to draw what they thought happened on the friend icon click, drawing the view on another post-it.
5. Then, the mission of the YoungDrive app was described: "If you would be given a quiz app now, what would you like it to be like? It's main purpose, is to assess what you've learned throughout the day. How would you design it? What happens when you open the app?". They were then divided into three teams, having limited time to draw the best imaginable YoungDrive coach quiz app they could. First, they designed the app from the top of their heads. They then pitched their results to each other.
6. On the next iteration, they were to suggest and design improvements how the app should be designed to improve learning, not only assessment. "It's secondary purpose, is to practise what you've learned throughout the day. What would be different?" They then again pitched their results to each other.

In addition to helping to understand the coaches, the design proposals were photographed and served as guidance for modifications of the existing app throughout the week.

Assessment via Quiz

At the end of each day, the app was used to test the coaches' knowledge. Each coach got either a smartphone, tablet or computer. The coach first took the quiz for the most recent session, and could then choose what to do next.

Figure 3.8: *Each person designed a YoungDrive app first for assessment, then adapting the app for learning. They were allowed to test the smartphones and tablets, and use the apps Quizoid or Duolingo for reference.*

As there was no back-end developed, Josefina by hand documented the scores of each coach, writing the name of the coach, the session, number of correct answers, and what questions had been answered wrong. Josefina then, when planning the next day, looked at the statistics, looking for trends that would inform the sessions for the following day. She also evaluated the quality of the questions, before creating the new question sets for the next day.

Experimenting with Quiz Before or After the Session

Since the coaches appreciated the app so much, we felt tempted to try what would happen with fun and learning if we tried using the app *before* a session instead of only after. During the rest of the week, we continued, finally finding preferences and tendencies from the coaches, via observation, interviews, and survey.

Experimenting with Design of Questions

During the week, extra tests were done to test the following:

- Number of questions per quiz
- Single-answer questions or multiple-answer questions
- Framing of questions
- Challenge level of questions
- Determining what made a question hard

- Having similar answers to increase difficulty
- Trying questions like "A goal should be X.", where the coach is filling in a blank via multiple-choice answers

Interviews with Josefina

At the end of each day an interview was held with Josefina, evaluating the findings from the activities. At the end of the week, a final interview was held. At the end of Day 5, a discussion was made with Josefina what it would look like to not record the answers manually, but pushing the results online. A co-creation workshop was held, where she drew an Educator Dashboard.

Design Workshop #2 in Zambia

In the last day of training, another workshop was held, where the question was: "What do you need in order to feel self-confident for your session?". This question was informed from iteration 1. The coaches was asked to write down one need per post-it during limited time, then choosing the most important need, and then writing a motivation why that need was important to them. Finally, they were asked to sketch what an app would look like, that addressed this need to them.

Assessing Test Items According to Bloom's Revised Taxonomy

In a reflection meeting with the teacher, it was assessed in a smaller sample, if the questions had truly tested the learning objectives of the YoungDrive lessons according to Bloom's Revised Taxonomy. When this was found valuable (that many questions scored low on Bloom's Revised Taxonomy), the supervisors from Linköping University were asked to control if the preliminary assessment of Bloom's Revised Taxonomy done by YoungDrive and the current author was correct, and to investigate all of the topic quizzes, week 1-10. The data and comments from the supervisors were shared with the current author, who presented the conclusions and insights needed to take into consideration for iteration 3, which led to the creation of a new question set to be tested with the Uganda coaches.

3.4.5 Iteration 3: Uganda Formative Test

Because of the many research and functionality needs, the study design of Iteration 3 became very important. A lot of development and ideation needed to be done. Iteration 3 had an even more detailed scope. Since the app now succeeds with the first use case, the coach training, now the focus could be on "learning at distance".

Trigger Material Used

It was chosen that "Are you sure?" + Improve would be included in the hi-fi material of the app, a flashcard approach would be tested as a low-fidelity material,

and to "record answer via voice" could only be presented as an idea during a field interview (experts said there would be usability issues, and the first-time smartphone user agreed). The Gold/Silver/Bronze reward system was included in the high-fidelity material.

Interaction Activities in Uganda

Instead of only testing the app in Tororo, because of the major changes, a test was held in Kampala, to early get feedback from an entrepreneurship student. During the test in Tororo, as Plan International staff are not allowed to support visiting coaches in the field during local elections, the co-project leaders in Tororo were consulted to carry out the field trips, so that it was still possible to attend the youth group meetings.

For the interactions, a big app test was held, a group interview was held, and then they were divided into co-creation workshop groups, with a presentation in the end. There was another partner meeting, with Plan International and Community Vision present. There was an app test with all of the coaches, "Testing the YoungDrive coach app", followed up by splitting into six workshop groups based on solving different problems discovered during the test.

The following day, there were three field visits to CBTs, observing how they prepared themselves for a youth session, and then observe usage of the app immediately after having prepared a youth session for assessing and becoming prepared for a session.

The app test simulated the app being used to assess preparedness for a youth session.

After the high-fidelity app test, it was tested with a low-fidelity prototype that the coach thinks aloud about the question, *before* receiving the multiple-choice answers. This test was done as a live quiz, and if the interviewee could not answer the question directly, the audience were asked and tested if they knew the answer (raised hands), and if nobody knew the answer, it was tested which of the multiple-choice alternatives they found most likely. During the afternoon, the coaches were divided into five groups focusing on improving the app experience for the coaches.

On Wednesday, the coaches from the field visits were gathered for a workshop. The purpose was to see how they acted when given the challenge: "Get 100% correct answers in one go, on the hardest quiz". A co-creation workshop ("Educator Dashboard") was held in parallel, with 3 CBTs and 1 project leader respectively.

3.4.6 Iteration 4: Uganda Summative Test

The focus of iteration 4 was a summative test, to assess the learning effect of the app. First, a pre-test was carried out in paper, including questions about the coach and an entrepreneurship quiz, based on a well-known study (Walstad and Kourilsky, 1998), see Appendix B. During the test, this was the first time that the app could send data to the server. Data was sent whenever a quiz was started, and whenever a quiz was finished. The group was divided into two, the ones who

brought manuals and those who did not. Those that had brought manuals, could use these with the app, see figure 3.9.

Figure 3.9: *Coaches answering the app questions for topic quiz 3 on Financial Literacy, and the coach guide quiz 9 on Action Plan.*

After the test, every coach was randomly divided into one of three groups. These groups of coaches were interviewed by the current author and the two Tororo project leaders respectively. In the coach groups, they were asked:

1. Why do you think you were correct or incorrect?
2. Do they like the app?
3. Are you stimulated by the app?
4. What did you like?
5. What did you not like?
6. When do you want to use the app?
7. When are you not able to use the app?

Their answers were recorded by the interviewee on paper as notes, to later be more deeply explained and discussed together with the other interviewers (these summaries were audio recorded for future reference). The exact answers were separated into three different mind maps, depending if they regarded learning, interaction design and service design.

3.5 Data Analysis Framework

The results from each iteration needed to be analysed, see the previous figure **??**. Depending on the kind of activity and results gathered, different data is collected. For qualitative methods, the output is almost always notes and observations, which are then concluded into insights. These insights, can then sometimes be used in the below analysis methods. For quantitative data, the output is almost always *quiz results* or data about the coaches. Before analysis, how this data has been processed is clearly described. Then, in the following sections, each data analysis methods is described.

3.5.1 Data Analysis

This section first presents relevant methods for analysing qualitative data, finally presenting analysing quantitative data. For analysing qualitative data in the research phase, Personas or Need groups are methods to analyse the though users of a product or a service. A Customer Journey Map supports understanding of how such users interact with a product or a service. Getting feedback from users involves getting ideas or suggestions for improvement of a product or a service. To analyse these, a sprint backlog is used to keep track of the priority of such feedback. If the features are then successfully implemented in the eyes of the users and stakeholders, can be tested using a sprint demo, where feedback is gathered for future work. These methods can be more closely read about below.

Persona or Need groups

When creating a product or a service, it is important to understand who you are designing for. Since the intended users might not always be around during the development phase, it helps to have a clear mental image of the user. Fictional examples of users are one method to do this, either by using personas or need groups. In this project, this is especially important, as a lot of development is offsite from the intended users, and the current author has no previous experience interacting with the intended users.

A *persona* is a fictional character, created to give an example of the user who you are designing your service for (Stickdorn and Schneider, 2010). Depending on how broad the target group is for the service you are building, the persona might have several different needs. Then, dividing the intended user groups in terms of designing for their different needs than their character traits might be more helpful (Widmark and Nissar, 2016). Dividing users by needs, if called forming *need groups*. A need group (like "The beginner" or "The planner") can be described by their behaviour and their need. Personas and need groups should be developed from research insights gathered from interviews or workshops with users and stakeholders (Stickdorn and Schneider, 2010).

Customer Journey Map

A customer journey map is said to provide a vivid but structured visualisation of a service user's experience (Stickdorn and Schneider, 2010). A typical customer journey is multi-channel and time-based, see figure 3.10. The customer journey map is beneficial both for understanding a user's touchpoints with a service, and also for collecting and analysing stories which explain why the journeys happened as they did (Stickdorn and Schneider, 2010). In this project, the Customer Journey Map is used to understand the activities that happens before, during and after having a youth session, and how these activities differentiate by different types of coaches (by persona and need group).

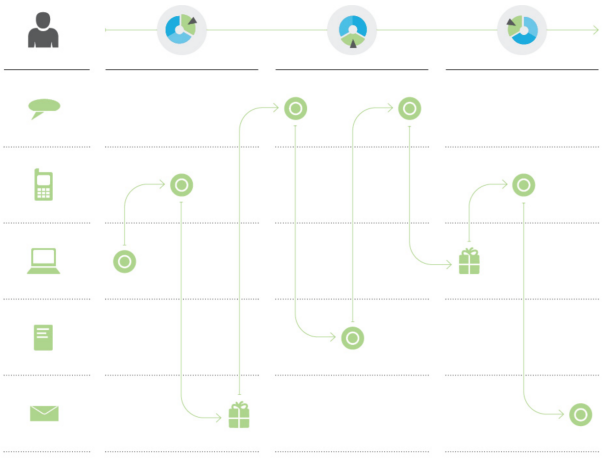

Figure 3.10: *General example of a customer service map. The map is divided by channel (could also be by a persona) and time (in this case before, during and after using a service).*

Sprint Backlog and Sprint Planning

During product development, to do a *sprint backlog* serves as a type of data analysis. When a high altitude of ideas and feedback has been gathered, a sprint backlog is a list sorted with the most important and urgent items on top. To work through ideas in a structured way, requests are categorized into what is called "stories". A good story answers "As a [user type] I want to be able to do [feature] so that [benefit]". Each story can then be broken down into tasks ("What needs to be done in order to satisfy the story?"). During the *sprint planning*, at the start of each iteration, the most important stories and tasks are identified. By analysing and then working on the most prioritized story and task, you ensure that you are always working on what is most important for the success of the project. In this project, the sprint backlog is used to determine what actions needs to be prioritized in order to accomplish the goals set up during sprint planning.

At the end of an iteration, as much as possible of the sprint backlog has been taken from "Not done" to "Done". The things not done, are either moved into a product backlog (to later be added into an upcoming sprint backlog) or removed (they were not important enough).

Sprint Demo

A sprint demo is an effective way to analyse the product created after an iteration, from the perspective of the users and the stakeholders. Kniberg (2005) says that a well-executed sprint demo attracts vital feedback from stakeholders, and ensures that tasks are 100% done. A sprint demo does not need to be complicated: the product is shown and tested with users and stakeholders, getting feedback for the next iteration. But without it, it would not be possible to know if the work done has satisfied the needs of the end users, or what the stakeholders thinks is important for future work. The sprint demo in this project is evaluated both with stakeholders and the end target group, the coaches. Evaluation is done by the interaction design principles: desirability (does the coach desire to use the product?), utility (does it meet the needs of the coach?), usability (is the coach able to use it without frustration?) and pleasurability (is the coach satisfied using the product?) (Clatworthy, 2010).

Analysing Notes

After an interaction, the notes (from interviews, workshops, tests and field visits) are transcribed from paper into the computer. This provides an opportunity to analyse the notes, think about their meaning and impact for the project, and categorize and separate notes that are comments from interviews (like a quote from a coach, which can help strengthen the description of a persona or a need group, or a feature suggestion that after analysis can be put into the product backlog) and insights from the current author, which can guide future development.

A *mind-map* is usable for sorting and categorizing notes that are thoughts. In this project the interactive tool MindMeister (Labs, 2016) is used. The tool has been used to sort through all interview notes and transcripts. Such an example include the evaluation interviews from iteration 4, that were divided into three different mind-maps: one fore learning, one for interaction design, and one for service design.

Analysing Log Data

Log data are often noted on paper during workshops and app tests. This includes the start time of a workshop, when and why a certain bug with the app appeared, or which coach chose to do an action in the app in a certain way. When these logs can help explain quiz results, they are often converted into comments in the Google Sheets where the quiz results are collected. This explains for example why iteration 4 has a coach that completed the coach quiz for week 9 for the summative test, but there are certain blanks in the Google Sheet: as the app did

not record the expected data, the known log data was inserted into the Google Sheet.

Analysing Test Items According to Bloom's Revised Taxonomy

Using Bloom's Revised Taxonomy, the question from each quiz in the app can be evaluated by it's position on the knowledge dimension and the cognitive process dimension. Each question is given a "score", consisting of a letter and a number. The letter had the range of A-D for the knowledge dimension (A. "Factual", B. "Conceptual", C. "Procedural" and D. "Metacognitive") and the number had a range of 1-6 for the cognitive process dimension (1. "Remember", 2. "Understand", 3. "Apply", 4. "Analyze", 5. "Evaluate" and 6. "Create"). The data and comments are brought together into conclusions, to be presented during a sprint demo. In this project, this Bloom analysis is essential for being able to determine the learning effect of the test items.

Analysing Quiz Results

Below, methods for analysing collected quiz results and pre-test data in the project are presented. The first method used is correlation, but if multiple variables need to taken into account, logistical regression can be used. As both of these are statistical tools, they rely on statistical significance in order to be trustworthy. In R programming language there are powerful tools for visualizing the correlation, for example using a "Correlation Heatmap", see figure 3.12. In small-scale app tests, such amounts of data might not be sufficient.

To find trends or oddities in a small data set, visualization techniques to discover the quantitative data by hand might be more beneficial. To enable a visualization, data needs to be acquired, enhanced, mapped to a geometry and lastly rendered (Ropinski, 2014). When rendered, a suitable rendering of a multiple-variable data set might be parallel coordinates, see figure 3.11. Made interactive, each axis can be made filterable by user interaction.

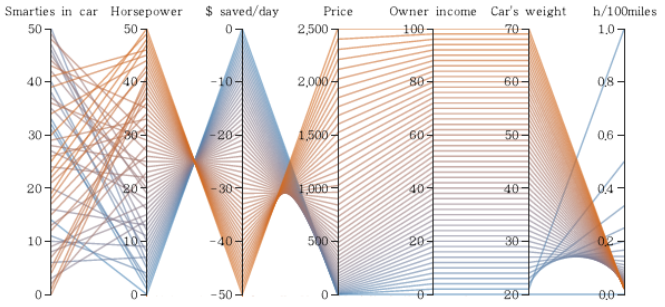

Figure 3.11: *Parallel coordinates enables you to visually observe relationships and trends between dimensions: positive or negative relationship (correlation or invert), or no relationship at all (random) (Lopez, 2016)*

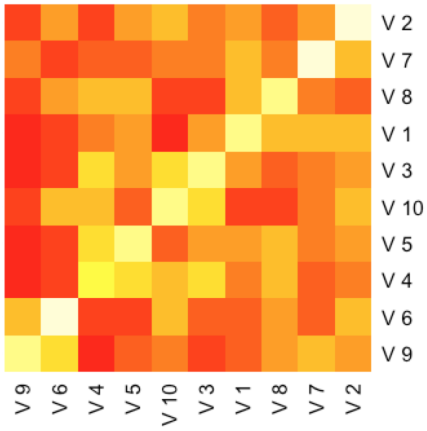

Figure 3.12: *A heatmap shows correlation axis against axis. The redder a segment is, the bigger is the correlation between two axises. In this example there is a high correlation between for example V4 and V9. To be statistically significant in correlation or linear correlation, a common measure is that the Pr value ("the p-value") needs to be higher than 0.05 (meaning a 95%+ probability).*

3.5.2 Quantitative Data Analysis of Quiz Results and Pre-Data

In this section, the steps needed to analyse the quantitative data is explained in detail, so that others can use a similar approach. In this example, quiz results are from iteration 4, but a similar approach has been taken with the manually recorded quiz results from iteration 2-3 as well.

Data analysis of quiz results were done first by a general overview in Google Sheets, by statistical analysis in R, and by a parallel coordinates visualization. The process to do this, is described below.

Step 1: Data Acquisition from Server

It was desired to store the data in Google Sheets, thus it was necessary to collect the MongoDB database content, and convert JSON format into a Google Sheets-readable format, like CSV. Multiple approaches were tried, and the Google Chrome extension called Magic Json (Team, 2016) was the one that worked without problems.

Step 2: Data Acquisition from Pre-Study

The Pre-study data acquisition was done by instead of looking at the paper-submitted pre-study evaluation forms, using the data processed into Google Sheets.

Step 3: Data Enhancement of Server Results

This section presents how data from the server was processed, to enable visualization mapping. To make the data easier to work with, the columns were reordered, and made sortable and filterable. Some columns were given conditional formatting, so it would be easier to spot irregularities, see figure 3.13. After this, some observations could be made.

To be able to compare the test results with the pre-test results, it was clear that it would not be viable to test every dimension against every dimension. Instead, since goals of the app evaluation had been predefined in the following way, the quiz results were summarized into a new sheet so that the following could be derived:

- % correct 1st try
- number of tries until 100%
- number of tries until 100% in 1 try

These could be calculated by having columns for:

- Quiz 3
  - Start time training
  - % correct 1st try
  - number of tries until 100% in 1 try

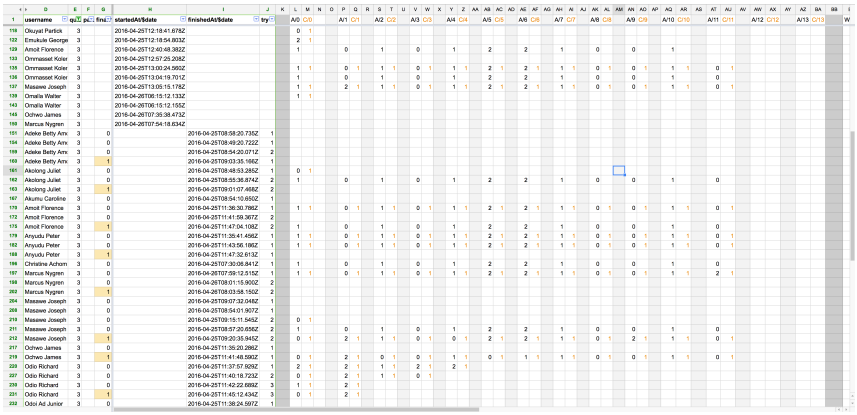

Figure 3.13: *The quiz results collected in a Google Sheet after data enhancement. In this figure, cell E1 has been filtered so that only results for quiz 3 are shown. Conditional formatting has been done with row G so that if it is a certification quiz, yellow color is used.*

  - Time difference start to end time certification

- Quiz 9
  - Start time training
  - % correct 1st try
  - Time difference start to end 1st try
  - Time difference start to passed training
  - Time difference 1st try to certified

Then, to see trends, color scales were again added. With ordinal values, a sequential color scheme is used (for example fastest time, from green to red), and with nominal values (like if they are female or male) where there is no right value, a qualitative color scheme is used. Now it was easier to spot outliers and trends, and giving validation to findings from notes and observations made during interviews, workshops and app tests.

Step 4: Date Enhancement of Pre-study Results

To see differences in answers more clearly, the data from the pre-study was made sortable and filterable. Then, the data was resampled for each column that had numerable (sortable) data in text instead of numbers, so for example "The day before" was changed to -1 and "The same day" to 0. In a similar way, school level was divided into four different groups, from 0 to 3, where 0 meant secondary, year unknown, 1 meant lower secondary, 2 meant upper secondary, and 3 meant tertiary.

After this, each column was given conditional formats using a color scale, using Google Sheets built-in functionality. This gave a visual way to quickly get a overview of the pre-test data.

Step 5: Data Enhancement by Joining Pre-test and Results Summary

The summary sheet and the pre-quiz sheet were joined, becoming a multiple-variate data set (several dimensions that were to bee compared with several other dimensions), see figure 4.18. A meeting with the university supervisors was held, so they could further give support in how to properly analyse the data. Since the two control groups showed similar means on the pre-quiz results, the two control groups were determined comparable.

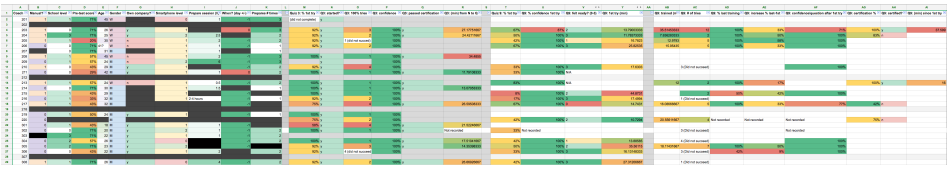

Figure 3.14: *The mutli-variate data set, made filterable and sortable in Google Sheets. Color scales and calculating means makes it easier to compare characteristics of the coach together with pre-test and quiz results. However, it is cumbersome and hard to quickly filter data on multiple parameters. For a zoomed in version, see* **??**.

.

To meet the challenges of using Google Sheets, a multivariate analyzation software or a visualization was suggested to discover the data in less time. It was hard to determine a suitable multivariate analysis software suitable when having so few data points. Principle Component Analysis or Cohen's kappa would not be suitable, neither was it believed applicable to do Linear correlation on all dimensions. After discussion with other Master thesis students working with analysing data from various disciplines, parallel coordinates was suggested. It would allow very quickly filtering of the data, finding correlations, and distinguish outliers and common characteristics.

To guide the usage of the parallel coordinates (as there is so much to discover in the data set), using R to do Logistic correlation was also done. A disadvantage with this method, is that to be statistically significant, many data points may be needed, and it was now known before-hand if the method would be useful. Probably, parallel coordinates would be the best method with analysing a small multi-variate data set.

Step 6: Visualization Mapping

The goal with visualization mapping is to generate renderable data, in this case for the parallel coordinates visualization. Thus, a new spreadsheet is added, specific for visualizing the data. Columns were deleted that would serve no visual

purpose (for example timestamps), gave all cells data values (even N/A when undefined), deleting users that did not have data, and shortened the column names so they would fit on the screen. The data was then exported from the Google Sheet into CSV.

Step 7: Rendering

For rendering, the JavaScript library D3.js, (Bostock, 2014) was chosen. It supports data-driven documents for visualizing data with HTML, SVG and CSS. It supports both JSON and CSV data. A visual framework for multidimensional detectives for D3.js was found called "Parcoords.js", Chang (2012). The example code from "Linking with a Data Table" provided the basis for the rendering. It allowed observing both the parallel coordinates visualization and the table data from the Google Sheet simultaneously.

To work, the example CSV file was replaced with the data from exporting the Google Sheets data. To benefit the visualization, also the colors were changed, and the toolkit's functionality to drag the axes titles around to reorder the dimensions was used, since the goal was to quickly compare and find correlations. The result is visible in 3.15.

3.6 Application Implementation

In this section, the prerequisites for the app is described, from the perspective of the user, stakeholders, and the developer.

3.6.1 User Needs

The technical constraints for the project, would need to affect the technologies used, if the project would be user-centered. On the client side, the app would need to be mobile and web based, consider non-access to internet, and not use a lot of battery, to work for the coaches of YoungDrive. That the app should be simple to use in this cultural setting leaded to design constraints and needs for evaluation.

3.6.2 Stakeholder Meeds

As the project was only three months, and the first month would be without digital development, time constraints were massive. However, to be able to answer how design affects learning, evaluation was needed to be done via data collection.

If no evaluation, there would be no need to write code, instead working with a low-fidelity prototype using pure design tools. Now, a data-driven approach was needed to measure, and therefore an app needed to be developed. On the server side, a database and API would be needed, to pull data from the database and push data from the client. Since internet was not always available, the client must be smart in its usage of pushing and pulling data. This would need to be investigated further into the project.

3.6.3 Choosing Frameworks for Creating the App

A JavaScript framework helps and speeds up the creation of building web apps. In the start of the project, Meteor (Inc., 2016) and Ionic Framework (Co., 2016) were both tested and compared with each other. It was decided that Meteor was the best way forward, partly because it would allow the app to be accessible on the web as well.

React (Facebook, 2016) (a JavaScript library for building user interfaces) was chosen as the front-end framework, having integration with Meteor and being relatively easy to learn and fast for development.

3.6.4 Creating a First Version of the App

Since Meteor was chosen, a multiple-choice quiz tutorial in Meteor was used to guide the first version of the app. Modifications were made, for example making it responsive and changing it to YoungDrive's graphic profile.

The app was pushed to GitHub, and first hosted on Meteor free storage, then available via youngdrive.meteorapp.com. For Android and iOS, it was made possible to install the app from the computer. For each day of the training in Zambia, new quizzes were added to the app, which created a belt path (see 2.4.2).

After iteration 2, a different hosting platform was needed when the Meteor free tier was removed, where Heroku was chosen. Staging environment using Heroku allowed changes on specific GitHub branches to deploy updates automatically on Heroku servers. The MongoDB database was created using the Heroku plugin MongoLab. A Meteor build-pack was used to allow Meteor to be used with Heroku.

It was also tested to upload the app to Android Play Store. The necessary steps from Cordova needed to be followed, screenshots needed to be uploaded, and some administrative tasks. After this, it only took a day for the app to appear on the Play Store, and everything worked satisfactory.

3.6.5 Enabling Data Collection

For iteration 3, as components grew, there was a need for a client-side router. The Meteor plugin Flow Router was used, as it was very popular with good integrations. For iteration 3, there was also a need to store data per individual, partly because the feature was prioritized from YoungDrive, but also because of the purpose of data collection. In order to store data per individual, a database and login would be needed. Because of technical difficulties, login and automatic data collection was not implemented until iteration 4, which is further described in the Discussion, section 5.1.3.

Login

To record data per user, would require login. This would be a usability issue for most problems, being first-time smartphone users. They need to find it intuitive,

user-friendly, and be able to remember the password in the future. A lot of different suggestions were through the ideation phase. The simplest login possible was chosen, after evaluation and discussion with experts: a 3-digit code, which was to be given to each coach during the test. Meteor had limitations with their auto-login module, which is very fast to implement. Thus, a 3-digit login was written manually. To summarize, the front-end was not problematic, however, implementing server-client communication so that it worked online and offline, was.

Online and Offline Database

If data was to be sent from the client to the server, there needs to be a database with Meteor Collections. An example app was made first, only using Meteor Collections. Meteor's use of Distributed Data Protocol (DDP), made app pushes feel immediate, even though data was not sent until there was Internet access.

However, it was found out that if it took more than 15 minutes to get online, the push would be aborted. For users that are seldom online, this would not be viable. An offline database was needed, and the plugin GroundDB was implemented. As it was cumbersome to get right, pushing the data whenever online, and hard to test (needed to wait 15 minutes each time), this was not ready for the interactions until Iteration 4. As a consequence, until iteration 4 of the app, no results were saved online via the app whatsoever.

For iteration #4, data collection was done by the app itself, which pushes data to server whenever online (it saves quiz start, and quiz finish). The server receives JSON data from the client, stored in the MongoDB database hosted on Heroku. Each data point is saved in a database called Results, with the signed in user (from the Users database). In the database, there are collections for Users, Quiz Lists, and Quiz Results.

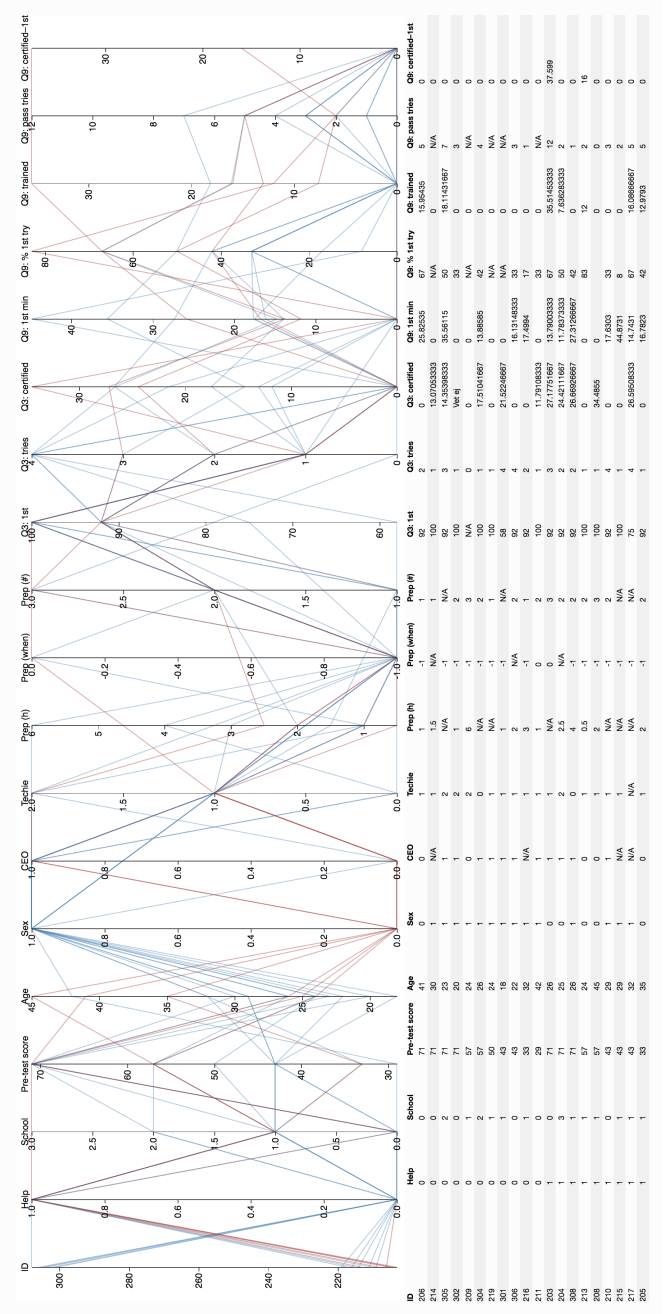

Figure 3.15: *The parallel coordinates visualization, done in d3.js. The visualization support draggable axes, filtering of data via dragging the sliders (which synchronizes with the data table), color assignment (like blue and red for men and women in the example), and hovering over a specific data point.*

4

Results

In this chapter the developed application is first demonstrated and explained, and how the app has evolved from iteration 1-4. Then, the results from all the qualitative and quantitative data collection in iteration 1-4 are explained.

4.1 The YoungDrive App: Iteration 1-4

After three months time, the app developed has been successful in addressing YoungDrive's goals with the app, with very high precision to the needs and context of the end users. In figure 4.1 the app development from iteration 1-4 are shown. In appendix A, all screens for iteration 4 can be seen in greater detail. The final app works on web or as an app, online or offline, on all of YoungDrive's Android and iOS devices. The app is fast to use, the back button on the phone can be used to go to the previous view, and the font size and images are consistent for each screen (which was not the case iteration 1-3, see figure 4.1 E-3. The goal was to provide a great learning experience, with a strong YoungDrive feel (embracing the values of fun, plus using the YoungDrive logo and colors).

Figure 4.1: *The app flow described as a timeline (A-I), per iteration (1-4). E.g. in iteration 4, login (B4) appears after choosing the quiz to take (A4). For a digital version where you can zoom in, see* `http://bit.ly/youngdriveapp-map`.

The quiz flow during iteration 1-2 was a standard multiple-choice quiz game, designed for assessment, but not for learning. In a scoreboard, the coaches could see which questions they were awarded points for and not, with a total score, see figure 4.1 H-1. In the end of iteration 2, whey were encouraged to go back to the start screen, to redo the same quiz again, or select a new one. 1 point is awarded for correct answers, whereas 0 points is given for incorrect answers.

In iteration 3, feedback for self-reflection was introduced. The coach answers Yes or No on "Are you sure?" for each question (figure 4.1 C-3), with the aim of increasing meta-cognitive skills, and being able to give personalized feedback in the score board.

Thanks to recording both if the coaches were correct and confident, the app can give very precise learning feedback (for example showing that the coach answered alternative B with confidence, but showing that A was actually the correct alternative).

In iteration 3-4, after observing their incorrect answers, and learning their correct ones (closely compare figure 4.1 D-3 and E-3), coaches were able to retake the wrong answers immediately.

In iteration 3, the score board was simply showing each correct question with the answer the coach provided versus the correct one. After giving feedback, the coach can train, and improve on incorrect answers and guesses, and when being 100% correct and confident, ideally take the whole quiz without faults.

For iteration 4, the score board is more personal, encouraging the coach to reflect on her own learning process. The feedback is designed to give a self-confidence boost ("You were correct, and you were sure" or "You guessed, but you were correct. You can answer with confidence the next time"), unlearn knowledge ("You were incorrect, but you were sure"), or encourage studying more when unsure and wrong ("You were incorrect, and you were not sure").

To discourage guessing during training, since iteration 3, the coach is shown number of tries on the quiz if they do not get 100% in their first try (see the top bar of figure 4.1 E-3). For iteration 4, the coach will also get a minus point (-1) for answering incorrectly but confidently "Are you sure?": "Yes". If uncertain, the coach should answer "Are yo sure?": "No", and no penalty will be given. The coach will later recieve feedback, either "You can be confident next time" (if correct but unsure) or "How can you get these correct next time?" (if unsure and incorrect).

The coach is encouraged to be well-prepared before trying again (either consulting the answers or the manuals). At 100% correct answers, they can get certified, by getting the whole quiz correct without faults. Failing a question on the certification would show that they have not *reliably* learned all the answers during the training. The more effort the coach has put into the training, the more likely the coach will be to pass the certification.

This certification quiz is the same as in training, but no longer do they need to answer "Are you sure?". If the coach can again get 100% correct answers, without any faults, they have proved they are correct and confident with the topic they should teach the youth.

If they fail a question, the score board will encourage more training and that they did not get certified, see figure 4.1 H-4. Since the coach did get 100% correct

in the training (maybe after a couple of tries) but not in the certification, the knowledge can be described as "Can do with effort". The coach can choose to take a new quiz, or take the same quiz from the beginning of the training.

For passing the certification on the 1st, 2nd or 3rd time the coach is awarded a Gold, Silver or Bronze medal respectively (after certificate try 3, no medal is given, only the certificate). Compare figure 4.1 I-3 to I-4: for iteration 4, the "certificate" mentions the persons name and the topic certified in, which is a form of gamification in addition to the intrisic motivation of rewarding deliberate practice.

In the following sections, you can read more about what data guided the development of the final app, as well as the results from the final app evaluation.

4.2 Iteration 1: Uganda Coach Visit

Here the results from the qualitative and quantitative data for iteration 1 are shown together with conclusions.

4.2.1 Interviews and Field Visits

Below, the most important results from the stakeholder and coach interviews are presented.

Stakeholder Interview Findings: Entrepreneurship Education Considerations

Through early interviews with YoungDrive staff, it is clear that YoungDrive's entrepreneurship education methodology goes hand in hand with the presented theory. It's mottoes are: "Dream big, start small", "Learning by doing" and "We have fun!" (Lönn, 2016). The YoungDrive program seems to be very appreciated by every party, especially the project partner, Plan International. The interviews with Plan International (see figure 4.2) showed great knowledge about YoungDrive's positive effect on the youth.

Both in regards to designing for the users and for the above reason, the app should be a complement to YoundDrive's existing training material and the structure of the program.

A challenging part of the work is that YoungDrive consists of both the practical skills of the entrepreneur, theoretical material of running a business, and an entrepreneurial mindset. Therefore, both how to assess knowledge, and build habits, needs to be examined.

As the stakeholder interviews answered "What's it like being a coach?", their perspectives could now be understood and summed into a early understanding of the coach situation. The stakeholder interviews heavily informed the Questionnaire guide, highlighting aspects that had previously not been taken into account.

Figure 4.2: *Meeting with Plan from the office in Tororo, making final preparations for the meetings.*

Coach Interview and Field Visit Findings: Teaching with Confidence

The purpose of the coach interviews and observations was to understand "What's it like being a coach?", from the coaches own perspective. Thanks to having similar interviews with stakeholders, the coaches' opinions and experiences could be compared. Also, more detailed answers on background, desires, experience and situation could be provided.

From the interviews and observations it was understood that CBTs can be responsible for from 7 up to 12 different youth groups in different programs, and such a high number places huge demands on the CBT. Even if there are only 7 groups, being behind on schedule or not being confident, can be very demanding.

The interviews and observations with CBTs, Project Leaders and stakeholders led to the realization that different coaches handles this differently well. All coaches possess high self-confidence in varying degrees in various situations, and as a result, quality among coaches is unbalanced, which stakeholders see as a challenge. Depending on the situation, everything is going according to plan, you are not confident, or you are falling behind with the schedule, you can be in one of these three Need groups:

- The ideal coach
- The realistic coach
- The challenged coach

It was discovered that coach confidence comes largely from being able to have good youth sessions, see figure 4.3. This is important, because according to the interviews, being a high-quality YoungDrive coach to a large extent means having high-quality youth sessions. For having a high-quality youth session, these are the most important attributes mentioned by the project leaders:

- Correct information
- Correct structure
- Time management
- Fun atmosphere

Figure 4.3: *Observing a youth session, the coach using the whiteboard to make certain concepts more clear.*

These findings were used to inform the Customer Journey Map workshop. In addition to the findings here, questions were also asked around ethnography.

4.2.2 Co-Creation Workshops

Two workshops were held, which together would inform future development of an application. These were the findings from those two workshops:

Customer Journey Map: "A Day As a Coach"

The three coaches understood the concept of the workshop surprisingly effortlessly, although the concept of postits and Customer Journey Map and Personas were unfamiliar to them. Contributing to the energy, was probably that the timeline and personas were largely informed by their interview answers. The workshop gave great insights for understanding the coach situation and the coaches themselves, both observing their behaviour during the workshop and learning about all the unknown activities involved in being a coach.

The first workshop was concluded with many important insights. Thanks to using Personas, it was discovered from the workshop that the difference between coaches and the quality of the youth training was more diverse than originally thought. This is especially true when the coaches prepares for their youth session, an activity which was regarded as important as the coach training, but where

quality of preparations were very divergent. This could be a great opportunity for delivering the app's promise of distance learning (see section 1.1). The teacher Josefina in an after-interview commented that while she can influence the coach training by her physical presence, she currently has no influence or insight into how the coach prepares, more than the co-project leaders' reports. When preparing for a session, there is a wide variety of *when* the planning happens, and how a coach judges what amount of preparation is enough.

Most coaches plan their next session during the morning, or immediately after a session with their group. Since a coach has somewhere between 7-10 groups (some even more), and the youth groups are at different modules, there is a lot of knowledge for the coach to handle - not only theoretical knowledge, but also the struggles of the youth, assignment presentations, workshops to be facilitated, etcetera. It is easy for a coach not to do everything as planned or as specified in the manual. Most of the coaches are said to be motivated by the possibility of becoming a better coach.

According to the workshop, most of the coaches assess if they were ready for a topic *after* a youth session. The feedback from the youth, as well as their questions and how well the coach can answer these, are the biggest informant. The exception is if the local project leaders comes to visit the youth session, but they do seldom have time to visit all of the coaches during a month, and during a month coaches should have taught 4 topics to all of their groups.

Smartphone Test using Quizoid and Duolingo

Quizoid (see figure 4.4) and Duolingo (see figure 4.5) were tested to understand the technical possibilities of the coaches, see figure 4.6. They had circa 45 minutes to spend in each app, with discussions after each app test. The result was that the app can place itself somewhere in the middle of the two, regarding difficulty level.

Figure 4.4: *Quizoid is a simple multiple-choice game (Apps, 2016), tested by coaches in iteration 1*

.

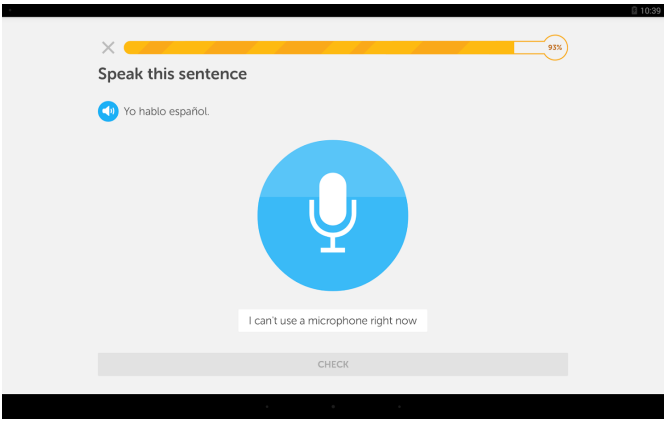

Figure 4.5: *Duolingo is a praised app for language learning (Duolingo, 2016), tested by coaches in iteration 1*

.

Figure 4.6: *Picture from the smartphone test, observing how the coaches comprehend using Duolingo and Quizoid for the first time.*

Two of the three YoungDrive coaches (two of them being the local project leaders) used a smartphone for the first time. The attitude towards using a smartphone was overwhelmingly positive. One of the coaches even mentioned: "Marcus, today one of my dreams have gone true" regarding using a smartphone for the first time. He even asked if he could borrow one of the devices during the remainder of the stay. Even for coaches that had never touched a smartphone before, some concepts were easily understood (like using the camera and Quizoid).

Other concepts were harder (accidently getting to the settings menu, unlocking the device, understanding advanced games, or training languages using Duolingo with advanced interactions). Point and click is easily understood, whereas sliding is much more unnatural.

4.2.3 Sprint Demo

During the evaluation meeting with Linköping University and YoungDrive, it was determined that Iteration #1 had provided preliminary answers for research question #1 (technical possibilities), #2 (contextual constraints), #4 (usability and learning), and partly #5 (how to use user feedback). Thus, the iteration could be considered very successful, and now, the development of the app could begin.

It is clear from the data that the motivation of the app should be to assess and strengthen the entrepreneurship knowledge and skills of the coach. For coach quality to improve was a desire from the stakeholders as well as the coaches themselves, even if they were also satisfied with the current results. This lead to a challenging situation how an app can address becoming a better-performing coach.

An app could increase accuracy of correct information. With an app, the coach could keep a record of the module content, and see when and if they do need to refresh their skills. It was discovered that a coach app can benefit not only the coach training, but also in a surprisingly precise way, what was called "distance learning" in section 1.1. Accountability if the coach is ready for a session by automatic assessment. A very important aspect to increase learning and confidence will be to give good feedback (see section 2.3.3).

With all the possible benefits of an app, it is definitely a problem that so few coaches have smartphones. Either continued development could be guided solely by the use case of having an app tailored for the coach training (where donated devices are available). But this would be to ignore that an app helping the coach to prepare for a session would be extremely beneficial, which discovered during the field visits to youth sessions and during interviews.

The motivation for using technology is very high, so one way forward would be that the app for distance training will reach only the users that can be given access to a smartphone, counting that more coaches will get smartphones in the future. Not using smartphones but feature phones (which all coaches possess), would mean building an SMS-based service (see 7.1.

As most coaches are already motivated to become a better coach and using technology, the advise from Sierra (2015) of designing for their compelling context can be followed. From a YoungDrive perspective, this might mean "Given a

teaching situation among the youth group, a great coach can teach an entrepreneurship topic more consistent with what the coach material said". Their performance in the YoungDrive app could translate into: "Given a question in the app, a great coach will get the right answer more often, and increasingly leverage the correct answer to their coach situation".

Next Iteration

Josefina agrees with the insight from the interviews, workshops and field visits, that the most important skill of a YoungDrive coach is having great youth sessions. It is a challenge that the coach surely needs to feel, but does not always possess, self-confidence for its youth session. This partly stems from the lack of practical experience being put into realistic situations during the coach training.

If self-confidence comes from being able to deliver Correct Information, Correct Structure, Time Management and Fun Atmosphere, an app strengthening these will surely improve youth session quality. According to Josefina, assessing and increasing Correct Information is the parameter she values the most highly.

It is agreed with Josefina that preparing for a youth session can have an increased focus. It is a worry that designing for both the coach training and preparing for a session might be too ambitious within the given time frame. If so, designing for the coach training is deemed more important.

4.3 Iteration 2: Zambia Coach Training

Here the results from the qualitative and quantitative data for iteration 2 are shown together with conclusions.

4.3.1 Co-Creation Workshops

There were two workshops held with the Zambia coaches: one at the beginning of the week (before being shown the app), and one at the end of the week (after having used the app for a week). Below, the learning from these two workshops are presented.

Co-Designing an App for the YoungDrive Training

The result gave an unbiased look at what the coaches expected from the app, what functionality wasn't important, and into their technical preferences. A simpler design than originally thought was deemed sufficient, and the simple sketches guided continued development of the app during the week. See figure 4.7, 4.8, 4.9 and 4.10 for reference. The time spent on designing the YoungDrive app in two iterations took the coaches circa 1 hour and 40 minutes.

From using the devices during the workshop (to find inspiration from other apps like Duolingo and Quizoid), it was found that most coaches prefer using the tablet (5 for tablet, versus 2 for smartphone and 2 for computer). Both the designs and insights gained were used throughout the week to further improve

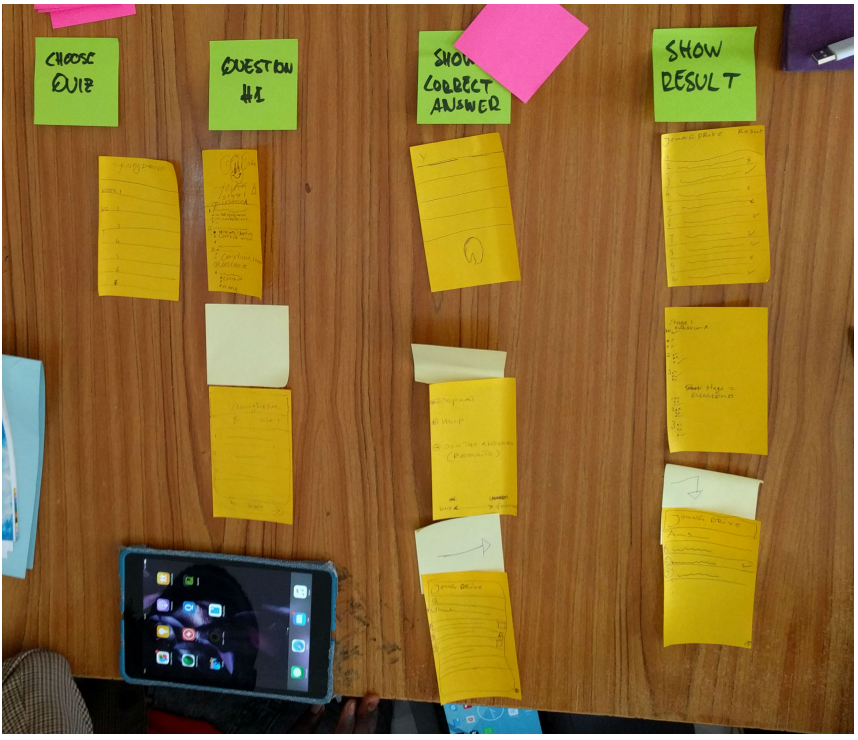

Figure 4.7: *The design proposal from team 1*

the simple app created at the end of iteration 1. The workshop gave great insights to who the coaches were and their thinking.

Understanding what Builds Confidence among Coaches

During this workshop, the focus was to examine what builds confidence among the coaches. Four themes were identified, after clustering the notes according to similarities: "I believe in myself" (3 coaches), "I believe in God" (2 coaches), "I am well prepared" (4 coaches) and "I am certified" (1 coach). See figure 4.11 for the different clusters. Josefina comments after the workshop: “I have a problem: There is no way I can control them how they have prepared themselves for a youth session.".

4.3.2 Quiz Results and Quiz Usage Observations

Here some of the quiz results are shown. This section also presents the results from the two workshops held, in form of lessons learned.

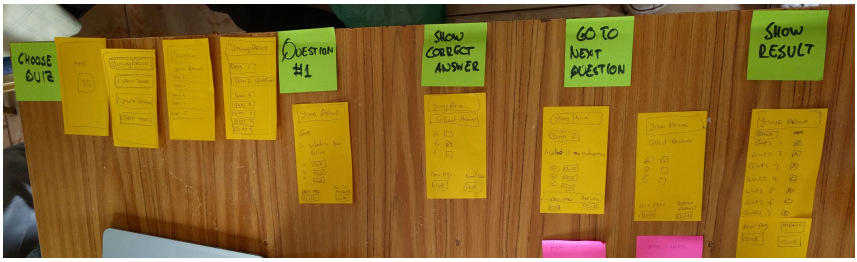

Figure 4.8: *The design proposal from team 2*

Quiz Results from the Coaches

See figure 4.12 for all quiz results, and figure 4.13 for a summary. Quiz results ranges from the worst on 24% (getting 4/17 correct answers) to 100% (which has happened on 52/101 instances, for all of the quizzes). [1] The quiz for the week 9 topic, "Goal Setting and Action Plan", was undoubtedly the most difficult quiz (17 questions, average being 76.22% correct answers, 18 quiz results submitted). The easiest quizzes were week 3 "Financial literacy" (11 questions), week 6 "Mastering sales" (9 questions) and week 7 "Sales day" (3 questions), where all of the results are 100%.

[1] There were one quiz for each of the 10 weeks of the youth training, one topic per week. In addition to these 10, there were also three bonus quizzes, regarding being a coach.

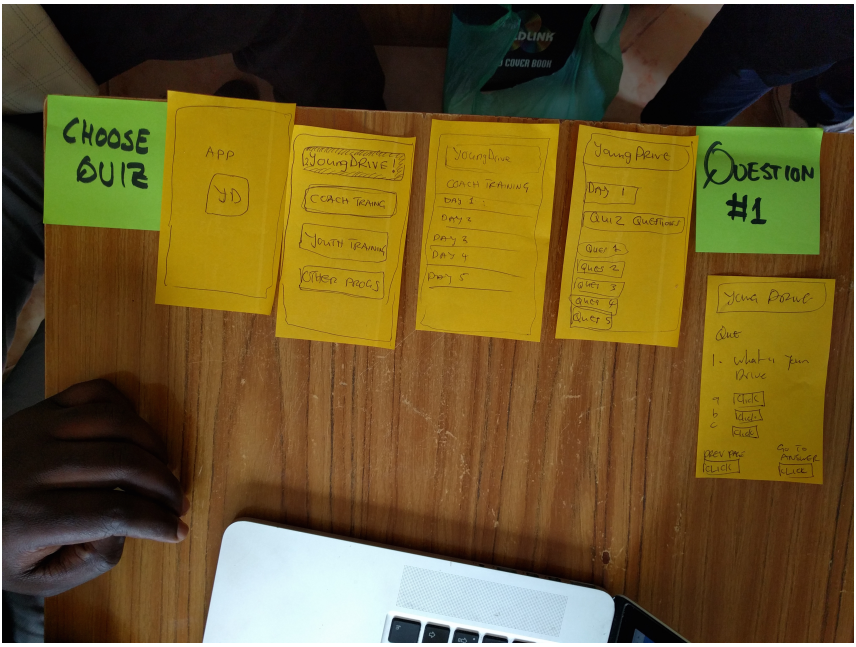

Figure 4.9: *The design proposal from team 3, part A*

Figure 4.10: *The design proposal from team 3, part B*

Figure 4.11: *The clustered postits with needs, together with included design proposals.*

	A	B	C	D	E	F	G	H	I
1	Day	Web?	Quiz	Pre/post-test?	T	C (%)	Time to complete	Coach	Wrong answers
2	5	y	9	1	19	N/A	0:17	404	N/A
3	5	y	9	1	19	100%	0:12	405	
4	5	y	9	1	19	100%	0:10	406	
5	5	y	9	1	19	100%	0:18	407	
6	5	y	9	1	19	95%	0:14	408	5
7	5	y	9	1	19	89%	0:14	403	8,19
8	5	y	9	1	19	74%	0:13	410	5,6,8,16,17
9	5	y	9	1	19	58%	0:14	401	5,6,9,10,17,18,6
10	5	y	9	1	19	37%	0:26	402	2,3,6,8,9,10,11,13,16,17,18,19
11	5	y	10	1	7	100%	0:11	405	
12	5	y	10	1	7	100%	0:17	407	
13	5	y	10	1	7	100%	0:16	408	
14	5	y	10	1	7	86%	0:05	405	5
15	5	y	10	1	7	86%	0:05	406	1
16	5	y	10	1	7	86%	0:20	401 & 404	5
17	5	y	10	1	7	71%	0:13	410	5,6
18	5	y	10	1	7	71%	0:21	403	5,6
19	5	y	10	1	7	71%	0:13	404	5,7
20	5	y	10	1	7	57%	0:21	401	3,5,7
21	5	y	12	1	10	90%	0:20	406	1
22	5	y	12	1	10	90%	0:06	408	10
23	5	y	12	1	10	80%	0:12	405	1,4
24	5	y	13	1	10	80%	0:10	410	1,10
25	4	y	8	1	10	100%	0:07	410	
26	4	y	8	1	10	100%	0:11	401 & 402 & 404	
27	4	y	8	1	10	100%	0:02	405	
28	4	y	8	1	10	100%	0:06	408	
29	4	y	8	1	10	100%	0:03	409	
30	4	y	8	1	10	90%	0:06	403	7
31	4	y	8	1	10	90%	0:04	406	7
32	4	y	8	1	10	90%	0:11	407	7
33	4	y	9	-1	17	100%	0:09	405	
34	4	y	9	-1	17	100%	0:06	406	
35	4	y	9	-1	17	88%	0:18	406 & 408	2,9
36	4	y	9	-1	17	88%	0:08	409	9,17
37	4	y	9	-1	17	71%	0:18	407	1,3,4,8,13
38	4	y	9	-1	17	71%	0:12	409	4,5,6,14,17
39	4	y	9	-1	17	65%	0:13	402 & 403	1,4,5,8,16,17
40	4	y	9	-1	17	59%	0:20	410	1,3,4,5,11,15,17
41	4	y	9	-1	17	53%	0:33	?	1,3,4,5,6,8,16,17
42	4	y	9	-1	17	24%	0:30	401 & 404	1,2,3,5,6,8,9,11,12,14,15,16,17
43	4	y	10	-1	5	80%	0:06	405	3
44	4	y	10	-1	5	80%	0:05	406	1
45	3	y	5	1	7	100%	0:09	410	
46	3	y	5	1	7	100%	0:07	405	
47	3	y	5	1	7	100%	0:08	403	
48	3	y	5	1	7	100%	0:13	406	
49	3	y	5	1	7	100%	0:13	407	
50	3	y	5	1	7	86%	0:11	408	3
51	3	y	5	1	7	86%	0:09	409	6
52	3	y	5	1	7	86%	0:16	404	6
53	3	y	5	1	7	71%	0:07	401	1,2
54	3	y	5	1	7	43%	0:20	402	2,3,4,5
55	3	y	6	1	9	100%	0:07	410	
56	3	y	6	1	9	100%	0:10	405	
57	3	y	6	1	9	100%	0:17	401	
58	3	y	6	1	9	100%	0:07	403	
59	3	y	6	1	9	100%	0:07	406	
60	3	y	6	1	9	100%	0:13	407	
61	3	y	6	1	9	100%	0:07	408	
62	3	y	6	1	9	100%	0:05	409	
63	3	y	7	1	3	100%	0:03	410	
64	3	y	7	1	3	100%	0:02	405	
65	3	y	7	1	3	100%	0:03	403	
66	3	y	7	1	3	100%	0:02	408	
67	3	y	7	1	3	100%	0:03	409	
68	3	y	8	1	10	100%	0:09	405	
69	3	y	8	1	10	100%	0:05	405	
70	3	y	8	1	10	90%	0:09	406	7
71	3	y	8	1	10	90%	0:05	409	5
72	3	y	8	1	10	80%	0:10	403	5,7
73	3	y	8	1	10	80%	0:08	408	7,8
74	2	n	2	1	14	100%	0:05	409	
75	2	n	2	1	14	86%	0:10	401 & 403	2,10
76	2	n	3	1	11	100%	0:14	410	
77	2	n	3	1	11	100%	0:04	405	
78	2	n	3	1	11	100%	0:04	402 & 405	
79	2	n	3	1	11	100%	0:07	403	
80	2	n	3	1	11	100%	0:12	406	
81	2	n	3	1	11	100%	0:29	408	3
82	2	n	3	1	11	100%	0:05	409	
83	2	n	3	1	11	100%	0:06	404	
84	2	n	4	1	6	100%	0:07	406	
85	2	n	4	1	6	100%	0:19	407	
86	2	n	4	1	6	100%	0:03	409	
87	2	n	4	1	6	100%	0:03	401 & 409	
88	2	n	4	1	6	100%	0:19	404	
89	2	n	4	1	6	83%	0:06	410	3
90	2	n	11	-1	10	100%	0:06	401 & 403	
91	2	n	11	1	10	100%	0:07	409	
92	2	n	11	1	10	100%	0:02	409	
93	2	n	11	-1	10	90%	0:14	402 & 405	10
94	2	n	11	-1	10	90%	0:08	406	10
95	2	n	11	-1	10	90%	0:03	409	10
96	2	n	11	-1	10	80%	0:13	407	1,9
97	2	n	11	-1	10	70%	0:11	404	2,9,10
98	1	n	1	1	?	?	0:16	405	?
99	1	n	1	1	9	?	?	408	?
100	1	n	1	1	10	?	?	404	
101	1	n	1	1	10	100%	?	406	
102	1	n	1	1	10	90%	0:12	410	1
103	1	n	1	1	10	90%	?	402	
104	1	n	1	1	10	90%	?	407	6
105	1	n	2	1	14	100%	0:22	406	
106	1	n	2	1	14	86%	0:03	403	2,10

Figure 4.12: *All quiz results from the Zambia coaches*

There are some amount of coaches that have taken the same quizzes multiple times. From this, interesting conclusions can be drawn. Having taken the lecture before taking the quiz shows a 15.0% average increase in quiz results, compared to an 12.8% increase with simply taking the quiz two times. This shows that a lecture is still more effective than learning via the app.

Time to finish a quiz took between 2 minutes to 33 minutes (3/4 quizzes that took longer than 25 minutes were on quiz 9, there is a correlation with number of questions). Most of the quizzes took under 5 minutes to complete, and results are always over 80% for these instances. This shows that the coach had a high confidence with the answers. Similarly, all of the scores under 60% has taken more than 20 minutes. This can be explained by that the coach is unfamiliar with the app or smartphone, and that the coach is uncertain of the answers.

It is seen in the quiz results that the quizzes did get gradually more challenging, as the average of quiz scores gets lower after day 3 of training, when it was discovered was welcoming the quizzes to get harder.

The quiz scores when quizzes are performed in group is very varying (from 24% to 100%), and it is observed that influence of another coach can lead both to a worse score than their individual average, and a higher score. It is interesting to see that the three coaches that have the worst average (67, 72 and 85%) are also the ones that have taken quizzes the least times individually, and the most taking quizzes together with others. This may show that the app is more effective when used individually, and there seem no other connection to previous entrepreneurial activity or background.

There are few correlations that can be made between the coach's background and experience, compared with quiz results or app behaviour, see again figure 4.13. Nothing in the coach information distinguishes the worst 3 performers in the quiz. For the opposite, there is however a noticeable behaviour that being confident, having trained youth, leadership experience, business experience, care for community, care for oneself and confidence to take on many youth, almost guarantees a high quiz average (which demonstrates learning) and quizzes done (which demonstrates motivation).

B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
Coach ID	Quiz average	Quizzes done	Quiz increase	Age	# of bads	Distance (min)	Confident or Shy	Trained Youth?	Been a Leader?	CEO experience?	Care for community	Care for oneself	# Youth/Group	# Groups	# = Youth
	Average:			26.8		33	70%	80%	80%	90%	50%	90%	19.8	3	59.8
402	57%	3		28	1	15	1	1	1	1	0	1	17.5	2.5	43.75
401	72%	4		26	2.5	20	1	0	0.5	1	1	1	8	1	8
404	85%	5		39	2	10	0	1	1	1	1	0	25	2	50
410	87%	11	15	22	1	90	0	1	1	1	1	1	25	3	75
403	91%	9	10	27	2	15	1	1	0	1	0	1	22.5	4.5	101.25
407	92%	9	29	22	4	45	0	0	1	0	0	1	10	4	40
408	94%	9	20	21	0	60	1	1	1	1	1	1	22.5	4	90
409	94%	13	11.66666667	21	1	40	1	1	1	1	0	1	12.5	4	50
406	94%	13	4	32	1	20	1	1	1	1	0	1	30	3	90
405	96%	14	10	30	0	15	1	1	1	1	1	1	25	2	50

Figure 4.13: *Quiz results summary compared with data about a coach*

For coaches with 0-1 negative remarks during the interviews (n=6) compared to coaches with 2-4 negative remarks (n=4), their quiz average is but 2% higher (87% to 85%). This is not significant, but this number is increased to 93% (an 8% increase) if the outlier from the group is removed (a coach which only did three quizzes with average of 57%). In the future, more research could be done into comparing coach data with quiz results.

Lessons Learned from App Test Observations

Below, lessons learned from the app tests are explained in regards to two categories: first by motivation aspects, and then by learning aspects.

Motivation Regarding motivation from the coaches, one coach wished the app to be available on the Google Play store immediately, so that "The app could be used on my spare time". Another coach, without a smartphone, said "I'll buy one", because the utility of the app seemed so high. There were also suggestions for improvements, like "The app should have notes, not only quesitions". Regarding usability, low resolution screens made the text be barely visible. This showed, that the app needed to be tested on a lot of different devices. This is particularly true, as on day 1, the coaches did not know how to zoom, which could cause accident refreshs, frustration or confusion. Even more importantly, the app needs to work offline! To be online on the phone is too expensive for the coaches, and too unreliable to give a satisfactory experience. Also, during testing, relying on internet can cause a lot of problems, especially if the teacher is alone.

When asked about what they thought about doing one quiz ("Graduation") as a pre-quiz (before the session), 10/10 said they liked doing the quiz before, and that it benefited their learning during the session. When asked why it helped, 10/10 said agreed on the statement "During session, it is easier to follow" and that "Giving the paper manuals before, scanning headings and pictures etc, would not help". So, using the quiz before the session increases learning, slightly decreases fun of the session, according to coaches. One of the coaches described it as "Fun and encouraging".

It was also tested to work in group or individually. The ones who answered, said that you learned more individually (3/3), and more fun doing it together (3/3). Doing it together, was enjoyable as it was "Very easy because of using different minds" and "We can collaborate to do better". It can be argued that the quiz being easier is not a valid motivation, but describes the learning in the app as a desirable difficulty.

When doing a post-quiz ("Goal setting") immediately afterwards versus at the end of the day (doing spaced versus massed learning), quotes were "I thought it was fun and challenging to do the quiz immediately afterwards", with another coach commenting "The mind was still fresh". After a discussion with the teacher, these were the results:

- 3/10 wants to do the quiz both before and after a session
- 1/10 wants to do the quiz before and at the end of the day
- 7/10 wants to do the quiz only immediately afterwards
- 10/10 wants to do the quiz immediately afterwards, and then again at the end of the day
- 7/10 wants to do the quiz immediately afterwards, and then a joined quiz with other topics at the end of the day

The high scores on using the app a lot indicates that they like the app. The teacher wants to listen to coach opinions, at the same time not spending more time than necessary on assessment.

Regarding motivation from the teacher, discussion with Josefina what would hinder her from using the app, she says: "Not doing data collection digitally works whenever they are 10 - but not with bigger numbers than that." Also, she does not wish the app to replace her. She enjoys teaching, thinks she has an important role, and suggests the app to be designed to support her and the coaches, not replace her. When asked about what technical difficulties appeared with the app, she acknowledges that bugs in the app was sometimes a hindrance to use the app effortlessly, and that a lot of testing (both high-dose, and high-scale) is very important to acknowledge these.

Learning Regarding assessing knowledge, coaches had surprisingly high quiz results, see figure **??**, and at day 3 they wished harder questions when asked. The response was to give harder questions at the other days, for example by introducing similar answers, and testing 4 alternatives instead of 3. The changes were appreciated. The app was later tested on a university student in Uganda after the Zambia training, both on early and later quizzes. The university student from Makarere University scored 100% correct, in spite of not having any entrepreneurship training. This showed that guessing was possible, or that the quizzes were too easy, a finding also shown by other tests with similar test subjects.

The teacher Josefina commented that this might not be a problem, as the YoungDrive coaches are not as skilled with using a process of elimination, and had indeed scored lower results on average with the later quizzes. When testing the app with refugee innovators during Humanitarian Innovation Jam in Uganda, similar results to the YoungDrive coaches were found. She explains this because people in rural areas are not being equally educated and skilled with reasoning as university students, the problem is not as big as could be. Josefina is very happy with the app, and reviewed the app in the following way to Plan International after the training:

"The (YoungDrive coach training) app is a great tool to measure how much the coaches learned and understood from the daily training; it provides a clear overview of what the coaches truly understood and what they actually still don't completely understand. Based on that information I as a tutor can adjust the training for the following day to make sure that the coaches understand everything correctly. The app also works as a motivator for the coaches; it´s clearly reflect their own daily performances. If they score high they become very happy and satisfied, if they score low they are eager to check their wrong answers.".

A coach scoring only 9/19 showed the relevancy of the quiz, as Josefina did not think she would have discovered that the coach was lagging behind otherwise. In the data, it was observable that the coach had done well together with others, but 3/7 when done individually. Josefina said about the 9/19: "This is where a control group would be beneficial". "He is often passive during open questions, but active during the team exercises."

According to Josefina, if you have a high score, you are ready. If not, you need

to redo the quiz. If you are 8/10 or lower, you are in the red zone. If lower than 10/10, they are not ready, the motivation being that what they don't know, they will teach in an improper way: affecting hundreds of youth. This is why Josefina thinks they should need all of the answers correct. Getting all of the answers correct can be supported by the literature, especially on deliberate practice where 95% reliability in the topic is needed for maximum effectiveness (Sierra, 2015).

Up until now, merely Correct Information has been assessed, not the other three factors. The fact that the app already is appreciated with assessing Correct Information, makes starting to assess the other factors interesting. Josefina informs that Correct Structure, Time Management and Fun Atmosphere would be the most viable to test *after* a youth session, not before. She notes, that *some* assessment could be made via the app before a session. This could to be further investigated.

4.3.3 Analysis of Questions in Regards to Bloom's Revised Taxonomy

Analysing the questions by each topic according to Bloom's Revised Taxonomy, showed that questions on average gradually reached higher levels on both the knowledge dimension and the cognitive dimension. In topic 1 of 10, most questions only reached the first level of the knowledge dimension ("Factual") and cognitive process dimension ("Remember"). In quizzes created at the end of the week (like topic 9 of 10), questions reached up to "Analyse" on the cognitive process dimension (level 4, not reaching "Evaluate" or "Create"), and the whole range of the knowledge dimensions ("Factual, "Conceptual", "Procedural" and "Metacognitive").

4.3.4 Sprint Demo

Below, the conclusions presented during the sprint demo with experts and partners is given. The quantitative and qualitative data from iteration 2 shows that the app works for "validating the coaches' level of knowledge during their education", which was the main purpose of the master thesis (see section 1.2. Now, the app should be tested in Uganda as well, and there can be an increased focus on "distance training" and "certify all staff").

Iteration 2 has translated the theoretical understanding of answering the research questions into practical experience. Now there are observable evidence for what the interactions from Iteration 1 showed:

- The purpose of the coach training should be to prepare the coach in having great youth sessions
- Therefore, this is what the quizzes should assess
- What it really means being a good YoungDrive coach, is having good youth sessions

Validating the coaches' level of knowledge The quiz results data shows that a lecture is still currently more valuable than taking a quiz in the app two times to improve (15.0% to 12.8%). For iteration 2, work on answering research question 4 has started. While the teacher has appreciated the multiple-choice assessment, challenges of designing test questions to support entrepreneurial learning has been found. It is clear that the app is valuable for assessment, increasing coach self-awareness and being a valuable indicator for the teacher. However, the questions formulated scores low on Bloom's Revised Taxonomy (Krathwohl, 2002) compared with YoungDrive's educational objectives for the topics.

There are previously found issues with using multiple-choice for assessment and learning, but they seem to become especially relevant in the context of teaching entrepreneurship. Either question formulation needs to be improved, or creative design solutions needs to be experimented with which can increase coach understanding and identify and reduce guessing. This needs to be further investigated for iteration 3.

Test with university student scored 100% correct, means that common sense can go a long way, and that the results can't be 100% trustworthy, and that multiple-choice questions has serious issues **??** - this, we already knew during and before the coach training - but it needs to be taken care of. Similarly business and leadership experience for the coach seems to lead to a higher quiz average, while a low quiz average can't seem to be connected to any of the coach characteristics found during the interviews. This makes it hard to design the app for different types of coaches, without testing other parameters, which should be done in iteration 4 for the summative test.

Distance Learning

In addition to the formative app tests, workshop # 2 heavily informs what is necessary when designing for use case 2, distance learning: preparing a session in regards to building confidence. The results from the workshop are somewhat surprising, factors not only those that relate to the four parameters from iteration 1 ("I am well prepared" and "I believe in myself"), but for some also "I believe in God" or "I am certified" (which relates to purpose 3 of the app). These should be considered for iteration 3.

Further, app tests expose how the app is currently not actively designed with learning in mind, and thus not distance learning. This is unfortunate, both because distance learning is important, and as the app test with refugee innovators shows that there is an opportunity doing entrepreneurship training in rural areas outside of YoungDrive's coverage area. In order for online coach-training to work for distance learning, learning and feedback, and not only assessment, is however essential.

While it may be technically possible, the teacher desires the app support her during the coach training, not replace her. Therefore, completely replacing the teacher with an app should be avoided. The teacher is very important for giving coaching and educating in a way that the app can't. But the teacher can also be empowered by the app. For the future, Josefina would have liked to be able to

stop coaches from having taught, if they do not have 90-100 % correct information on the subject. Today, Josefina can not assess this. This means that some coaches, are teaching incorrect information to hundreds of youth. Here, the quiz has a very good need to fill.

If wrong on an answer, the app today has no means of giving high pay-off tips to get to 100%, or exposing you to deliberate practice or perceptual exposure. If the coach gets 9/10 correct answers reliably, or gets 5/10 answers with guesses, the coach still needs to retake all answers, not having learned the correct answer before taking the quiz again.

How to develop the app to solve these issues, is not obvious. Multiple strategies could and should be used. The app could benefit from introducing smart feedback encouraging a growth mindset ("You did not get 100% *yet*") (Dweck, 2014).

Next Iteration

After the meeting with the partner and expert group, the following was concluded from iteration #2:

- The app is partly working on assessment now, but not for learning. Are coaches really learning via the app, especially learning to be better coaches?
  - Multiple-choice is flawed in its current form. How can guesses be identified and reduced in a multiple-choice format? How can answering questions improve confidence and encourage learning?
  - How can questions be formulated in a way that teaches entrepreneurship, which is so practical?
- The need for a field app still feels relevant (especially for sessions long since the coach training)
  - An app could be used, either before you start planning (to guide what you need to study the most on), or after you think you are ready (so you can assess and improve).
  - When designing the app, it is concluded that an app for coach training, and an app to use before a youth session, should be able to be the same app if possible, since the purpose of preparing the coach to be great with its youth session is the same.
  - Discussing the importance of self-reflection after a youth session with Josefina, led to asking more of such questions in coach quizzes. While Fun Atmosphere can be hard to assess using multiple-choice, can Correct Structure and Time Management be assessed?

After the partner and expert meeting, it was decided that the following needs to be done for iteration 3:

- Make sure that the coach actually learns the desired educational objectives, by performing the following:

  - Create a new quiz guided by Josefina, "Are you ready for Session 9?", also to test if Correct Structure and Time Management could be assessed using multiple-choice
  - See if design additions to multiple-choice can increase learning in-line with Bloom's revised taxonomy

- Design quiz app for learning, focus on field app, and have a design that works stand-alone from the YoungDrive coach training in mind.
  - Investigate the effect of giving growth mindset feedback in the app (The Power of Yet approach)

- Test if the app created in Zambia could work also in Uganda
  - This also means converting all the questions from the new (Zambia) manual to the old (Uganda) manual, since both structure and content of the manuals has changed.

4.4 Iteration 3: Uganda Formative Test

Here the results from the qualitative and quantitative data for iteration 3 are shown together with conclusions. For the coach training in iteration 1, designing the app only for assessment but not for improvement was acceptable, since Josefina could give feedback in person. This meant that scoring low (like 5/10) or high (like 9/10) on a quiz after a session, the coach not taking the quiz again was not mandatory. Josefina could observe the results, and complement the missing information in an upcoming lecture or give immediate feedback herself.

In contrast, image the situation of using the app being weeks or months *after* the training, and Josefina is no longer available. Instead, the coach uses the app exclusively one hour before a youth session, to assess if the coach is ready. If a coach receives the score 5/10 or 9/10, what should happen? Clearly, in this case the app should be designed to not leave the coach alone, but help the coach to get smart feedback, encourage the coach to try again and improve their score, and again assess if the coach is fully ready for teaching the youth in the given topic. This new design of the app was the focus of the interactions from iteration 3, the results observable below.

4.4.1 Initial Evaluation of the New Version of the App

Before going to Kampala, because of the major changes to the app, the concept was tested with an entrepreneurship student in Kampala and the Zambia teacher, Josefina. The two tests informed that the app was now ready to be tested with the coaches in Tororo.

The entrepreneurship student's overall opinion on the app was: "Can you give me the link, because I'd love to do more of this. I think it's amazing.". There were some issues found with phrasing: "Improve" should be renamed, because it is not intuitive what the button would do. The student was also surprised that

the certification did not include something substantial (meaning it felt hollow). The student would have preferred unlocking a business challenge (showing self-determination), or something where he could get a learning reward instead of a "well done" and a badge (showing the student was not motivated by achievement in itself), see figure 4.1 I-3). This test was very valuable, and gave early insight to how the Uganda coaches might act within the app.

The teacher in Zambia, Josefina, was consulted to comment on the app. When asked for an opinion, Josefina answered: "I like the idea that when the coaches have answered all of the questions correctly, they can consilidate the knowledge by the certification test, when the coach should get 100% correct on their first try." This verified the relevancy of the taken approach of separating Training and Certification, see figure 4.1.

4.4.2 App Test: Learning Gain and Effect of Functionality for Formative Assessment

There were 25 coaches active during the app test. The test clearly showed evidence between the difference between designing for Assessment and Learning. See figure 4.15 and figure 4.16 to understand the setting.

Before the quiz started, the coaches were asked to raise their hand if they felt proficient with using a smart phone. 8 out of 23 said yes. After using the app, 16 thought they were proficient (25% increase), while 5 said low proficiency, and 2 said no (we don't yet feel proficient, still fear).

The test was done in pairs, because of lack of devices. Manual data collection of quiz results was not viable with coaches more than 10 people (as Josefina said in Zambia), it got to hectic, which is why not all quiz results were recorded, only when coaches raised their hands, which they often did only when they got 100%, and some never raised their hand. The results gathered can be seen in figure 4.14. Finally, the live quiz after the app test, to evaluate the low-fidelity prototype of including open-ended questions, proved to be great for learning, and proved possible to be a great addition to the high-fidelity material.

A couple of notes can be made from the quantitative data. Comparing quiz results between different characteristics however, some conclusions can be made. 6 groups consisted of two CBTs, and 5 groups were mixed (1 CBT and 1 Youth Mentor in the group). Among the CBT groups, the average score on a quiz was 77%, while the presence of a Youth Mentor increased the average to 98%. The groups were free to choose whichever quiz they wanted to. Only one group chose to redo the same quiz (scoring 100% both times), while all others varied between different quizzes. Out of 18 quiz instances, 85% was the mean (versus a 89% average over 105 instances in Zambia), with 1 instance of 0% (0/15!), 35%, 53%, 71%, 2 instances of 83%, and 12 instances of 100%. The lower result than in Zambia could be that in Zambia the topics were still fresh, and that in Uganda, much of the information from the earlier topics of the training has not been repeated in a long time.

Regarding pre-knowledge and learning, only 2/18 were 100% correct on their first try, which shows the importance of feedback to improve. Only 2/12 gave up

before getting 100% on a quiz. For those that continued, 3 groups got 100% correct on their 2nd try, 3 on the 3rd try, 2 on the 4th try, and and 1 on the 6th try (meaning more coaches pass the quiz on earlier than later tries). It is difficult to see if there are major differences in quiz difficulty, as the sample of doing the individual quizzes are so low. On both the topic quiz and coach guide quiz for week 9, it can be seen that all groups eventually scored 100%, as well as 3/4 for week 8. This may be explained by that most of the coaches were carrying out week 8, 9 or 10 at the time of the app test.

	A	B	C	D	E	F	G	H
1	#	ID #1	ID #2	Quiz	C	T	%	Tries on quiz
2	1	220	207	4	5	6	83%	1
3	2	209	210	CG9	13	13	100%	1
4	3	205	212	7	8	15	53%	1
5	4	305	217	2	14	14	100%	3
6	5	201	213	9	3	3	100%	1
7	6	306	211	4	5	6	83%	1
8	7	220	207	2	10	14	71%	?
9	8	304	214	2	14	14	100%	4
10	9	305	217	8	15	15	100%	6
11	10	218	?	2	14	14	100%	2
12	11	205	212	CG9	13	13	100%	3
13	12	306	211	8	15	15	100%	3
14	13	209	210	CG9	13	13	100%	2
15	14	201	213	8	0	15	0%	2
16	15	206	203	2	5	14	36%	2
17	16	201	213	2	14	14	100%	2
18	17	302	204	9	?	?	N/A	4
19	18	301	219	8	15	15	100%	4
20	19	303	215	CG9	13	13	100%	3

Figure 4.14: *The data gathered from iteration 3. A is order of quiz submission, B and C is Coach ID for coach 1 and 2 in the group, C is quiz chosen (CS9 meaning "Coach guide week 9"), D is correct answers, E is number of questions, G is percentage correct answers, and F is recorded quiz try on the particular topic at the time of submission.*

Because the quizzes were carried out in pairs, the results are non-traceable to

individuals, and thus also to pre-knowledge about the coaches or the results for iteration 4. This means, it is hard to track how their results compare over time. However, the insights from this data can guide research for future workshops, interviews and app observations.

Figure 4.15: *Two coaches using a tablet for the formative app test. The coaches worked in pairs. After the app test, interviews was held, before co-creation workshops started.*

Observations during test-session

Regarding usability, the most negative thing from the app test, was that the app was not user friendly for first-time smartphone users. There were a lot of bugs, the most damaging for the app test being resizing of the font size for each new question, see figure 4.1 E-3. This forced some coaches to try to zoom on the devices, even if they did not know how. This could in turn cause refresh of the web page, and sometimes there was no Internet available. Thus, this can explain why some quiz tries were started but never finished.

This was the first time true frustration was shown. Out of 23 respondents, 7 rated the app easy, 11 medium, and 6 hard. This was not viable, in terms of reaching YoungDrive's goal with the app: "It should be easy to understand, pedagogical and enjoyable to use, and the coaches should think it is fun and meaningful to learn via the app", see section 1.1.2).

Another reason for the low usability, is that cognitive load seemed to be too high. The feedback was not scaffolded enough, so coaches did not have enough energy to assess all of their results carefully before taking the test again via "Improve". One user did not want to press "Improve" until having read the manual. The motivation was: "Not because that is what the info says, but because I can learn more from the manual, about more than what the questions says." This is

Figure 4.16: *A variation of smartphones and tablets were used. In one case, the battery died on one of the device so it needed to be replaced with a computer. It was the first time the coaches used a computer, and they learned quickly and eagerly.*

in fact the preferred behaviour from Josefina, and the app should continue to further encourage only using the app training or certification mode after having prepared via the manual. This way, the app is still assessment, but it is "learning by thinking", with feedback. In iteration 4, comparing those that are allowed to use the manuals with those that are not allowed to, would be interesting.

Regarding learning, four new ideas were tested during the app test, assessing the new pedagogical model of the app. Item 1-3 were tested in the app (the high-fidelity trigger material), see figure 4.1, and item 4 was tested via low-fidelity trigger material, asking a coach (and then the coach audience) if he or they could give the answers from memory without being shown alternatives.

1. "Try again"-button. When clicked, your wrong answers are repeated.
2. If 100% on the 1st try, gold. On 2nd try: silver. On 3rd try: bronze.
3. Ask meta-cognitive questions, "Are you sure?", at the end of each question.
4. Speak your answer to the question before you are shown alternatives.

Item 1, 2 and 3 were determined good after the interviews. Item 4 showed relevance, but implementation in the app would give many challenges (mainly how to assess if the coach could give the right answer without giving different alternatives. Voice recognition or free-form answers are hard to analyze, to implement, and to use by a first-time smartphone user).

4.4.3 Follow-Up Co-Creation Workshops

There were room for doing five co-creation workshops during iteration 3, which were done immediately after the big app test with the same 25 coaches. For the following workshops, the coaches themselves could propose topics. These were the five suggested topics by the coaches:

1. Finding the YoungDrive icon after unlocking the device

    - The outcome led to in iteration 4, only the YoungDrive icon being on the start screen, and no other apps.

2. Making the app more user-friendly

    - While the proposals was not very concrete, this led to a realization that the app needed to be more user-friendly, and thus gave a larger focus on this for iteration 4.

3. Finding local examples of entrepreneurs to inspire the youth

    - This lead to the realization that most coaches having a hard time finding local success examples of entrepreneurs. An app could address this need in the future, both having a bank of successful local entrepreneurs, and booking them for visiting a youth session.

4. How to get access to smartphones (costing no more than 70 USD)

    - The action was very concrete: a Plan International staff voluntarily participated in this workshop, they contacted two retailers of smartphones, and concluded that 1) coaches could utilize the youth saving group to afford buying their own devices or 2) Plan International could buy and then borrow devices to the coaches, coaches being prepared to pay if they got lost or damaged. The results showed that both stakeholders and coaches are very eager to be equipped with smartphones, seeing the benefits.

5. Becoming a better coach via other apps (like Google's products)

    - This workshop was very interesting, as coaches found numerous uses of a smartphone which would benefit them in their work. The coaches even figured out how to use Google voice to ask questions, like the most profitable company in Tororo, or getting directions, or using the app for translations. It is evident that equipping the coaches with smartphones has very concrete extra benefits other than the YoungDrive app, and shows a tendency that the coaches are more eager to use the smartphone for utility than for entertainment.

4.4.4 App Tests with Three Coaches Before and After a Youth Session

Below, the results from the app tests with three coaches before and after having a youth session are shown.

App Test with Coaches Before a Youth Session: Collection of Affective Reactions, and Proposals for Improvements

There were three coaches testing the app after having prepared for their youth session, but before carrying out the teaching. Some things were notable from the field visits:

- "Are you sure?" is understood intuitively (you can't progress without answering), but some coaches deliberately answer "Yes" even if they are not sure.
- Idea to highlight different words of similar answers, to increase speed
- In summary, if wrong, show the other alternatives either way, not only the wrong answer
- Idea for future work: "Go to participant manual" within the app
- If correct and unsure, she says "I still feel good". "Include it in wrong, because maybe I was still guessing".
- Change button to "Become certified", to increase likelihood to press the button. As of now, it was not obvious.

When she did get certified, she said "I feel good". When asked why, she said: "They have appreciated what I have done". The next day, the same three CBTs gathered at the Plan International office to do an app test on the hardest quiz.

Having a service mini-sprint after the field visit, quick iterations could be made to the app. One such example comes from the field visits. Originally, it was believed best to use Gold/Silver/Bronze in the Training mode, and "Are you sure?" in the Certification mode. User tests showed that the other way around was better, and this was changed for the next meeting with the coaches. This example shows the relevancy of testing the app with the intended users, as it had not been evident from the tests with the Kampala student or the teacher.

A service design approach was used, first observing how preparations was made without the app, and *then* introducing assessment via the app, followed by interview. What was the most valuable feedback from the field visits, was to see that the app had indeed been a perfect fit for use in the field before a youth session. However, it was not possible due to time limitations to follow the coach to their youth session afterwards, to see the actual effect of preparing via the app.

App Test with Coaches After the Youth Session: Quiz Results and Analysis

Several conclusions could be drawn the day after from a joined app test with the same three coaches from the field visits. Two of the coaches were competing with finishing the coach guide quiz 9 (see results in figure 4.17, the coach guide quiz for the topic they had taught the day before. The third coach (who arrived late) offered feedback on the app.

	A	B	C	D	E
1	Coach	Try #	%	Time (min)	Manual?
2	203	1	33%	13	1
3	203	2	N/A	10	1
4	203	3	N/A	10	1
5	203	4	25%	10	1
6	203	5	67%	10	1
7	203	6	100%	3	1
8	203	C1	67%	17	1
9	203	1	N/A	5	1
10	203	2	N/A	6	1
11	203	3	100%	6	1
12	203	C2	100%	12	1
13	204	1	33%	17	0
14	204	2	42%	8	0
15	204	3	33%	20	0
16	204	4	42%	20	0
17	204	5	100%	22	0
18	204	C1	75%	5	0

Figure 4.17: *Data analysis in Google Sheets from the quiz results during the workshop for the coach guide quiz 9: "Are you ready?". It is observed how the two coaches (id 203 and 204) scores for each new try, where C1 means "Certification try 1". N/A means that the quiz score were not logged at this specific time. "Manual?" means that if 1, the manuals could be consulted.*

The two coaches did 17 quiz tries on quiz 9. Coach 203 (referred to as "Beatrice" in the text) was allowed to use the manual, while coach 204 (referred to as "Rehema" in the text) could only consult the feedback via the app, see figure 4.1, D-3 and H-3.

Apart from the quantitative quiz results, some comments were made. Beatrice having 4/12 correct answers on the first question, when asked from the quiz question 13 "How comfortable and ready do you feel right now to carry out session 9? (There is no right answer, just be honest with yourself!)", answered " Ready, but I probably still want to look in the coach manual and participant manual.". The other alternatives were: "Not ready at all, I need to prepare myself more by using the coach manual and participant manual.", "Somehow ready, but I still need to look more in the coach manual and participant manual." or "Very ready and comfortable.".

Beatrice got 4/12 correct answers on her first try, but eventually she did pass the certification, getting 12/12 correct answers in one try in 102 minutes from the quiz start. It took her 43 minutes from her quiz try 1 to passing the training. She then failed the certification try 1 (i.e. not getting Gold) after spending 17 minutes with it. Thus, she needed to go back to the training. Back in the training, now she got 8/12 instead of 4/12, and passed the whole training in 17 minutes. Then, she got 12/12 in the certification test (earning Silver), in only 12 minutes.

Rehema also got 4/12 correct answers on her first try. Not understanding the "Improve" button, she repeatedly went back to the home screen and retook the whole test (like in iteration 2). This was a slow approach. On her new tries, she got 5/12, 4/12, 5/12, which showed that learning was too hard. Similar results had been found in iteration 1 using Duolingo, where the app failed to train the coach to get a better score. After this try, she was explained the "Try again" button, got 7/7 (spending 22 minutes with the questions compared with her 16.25 minutes average doing all 12 questions, which showed she really put effort into analyzing the answers properly). Unlocking the certification test, she got 9/12 in 5 minutes, an improvement from 4/12 on her quiz try 1 75 minutes before.

The results from this quiz shows that:

- "Improve", only needing to repeat the questions you are not sure of, does improve learning quality and speed, compared to retaking the whole quiz again after each try - it makes learning more focused on what you need to train

- It needs to be clearer that you should press "Improve", for example by changing the text to "Try again"

- The time needed to become reliant on the session takes too long. To pass the rule of thumb for deliberate practice, a session getting to 95% reliability should take 45-90 minutes. For Beatrice, it took 102 minutes from quiz start to being 100% reliable. For Rehema, she did only manage to get 75% reliable within the same time frame. Either scaffolding needs to increase (dividing the quiz into smaller chunks), or learning effectiveness needs to increase (for example by better feedback).

The data and observations also shows that learning Correct Structure and Time Management via multiple-choice is not effective. Especially, this is shown by the time it takes to get a high score. To score well on such a test, the coach would retrieve from memory using a clear mental image. In its current design, getting a clear mental picture is not supported from the multiple-choice design. See Future Work in chapter 7 for a further discussion how this Correct Structure and Time Management can be better assessed in the future.

4.4.5 Sprint Demo

A sprint demo concluded the findings from iteration 3. Now the coaches could not only assess, but also *learn* Correct Information, which was successful, but needs to be done more effectively. It took an unacceptable amount of time to reach 100% proficiency on all the quizzes. This was especially evident, on the quiz on Correct Structure and Time Management, "Week 9: Are you ready?", when it took a coach 102 minutes to reach 100% without errors. In iteration 2, when "Improve" did not exist, it probably would have taken even longer.

For the first time both signs of learning via perceptual exposure (many questions during a limited time, by trial and error) and deliberate practice (via learning via reflection) could be identified from the app. It is just that the app as of now is quite inefficient, especially in terms of speed, so while the ideas are there, the criteria are not fulfilled.

The focus had been on "I am well prepared", but also including "I am certified.". It was shown that most coaches does not care about the simple gamification aspect of "I am certified" (which the workshop already had shown) but that they do care about their learning progress and learning results. The app could further embrace this.

If there is one thing additional learned during the iteration, it is the insight that data is knowledge, and knowledge is powerful. A realization is that both the developer, the coaches, the teacher, and the project partners can gain important insights from the data. Adding "Are you sure?" to each quiz question, coach understanding was amplified, because now, also the coach's attitude towards learning can be evaluated. See more about this in the Discussion chapter 5.

Next Iteration

To improve effectiveness for the next iteration, a couple of goals were chosen. While the app would work well for the Ugandian coach training, the use case of a youth session was not good enough yet. Mostly, this is in regards to that it takes too long time to improve via the app, and that the feedback is not sufficient. This leads to introducing these forthcoming goals, with the associated recommendations:

- Improve Deliberate practice. The criteria for Deliberate practice is not fulfilled today.

  - Follow the recommendation to design so that knowledge in a topic can go from unreliable to 95% reliability within one to three 45-90-minute sessions.
  - If this is not possible from changing the learning tactic, don't continue trying: split the quizzes into smaller pieces (Sierra, 2015).

- Improve Perceptual expose
  - Divide the learner's expertise according to Sierra (2015), "Can't do", "Can do with effort", and "Can do effortlessly".

- Increase the use of questions to prompt self-monitoring and self-evaluation (Sitzmann and Kanar, 2008).
  - Using "learning by thinking" and encouraging a growth mindset, can benefit reaching metacognitive skills on Bloom's Revised Taxonomy.
  - Help the coach to analyze and evaluate its own learning, possibly improving faster in the app.

- Improve feedback to reflect that the teacher does not want to encourage coaches to have their youth session before they are 100% confident with the material

- Data collection should be online and needs to be individual, so that the data is increasingly without faults and can be more easily analysed
  - To do it online means that there needs to be a database, but also a login, so individuals are traceable.

4.5 Iteration 4: Uganda Summative Test

Here the results from the qualitative and quantitative data for iteration 4 are shown together with conclusions. In these activities, there were 25 coaches present during both the app test and the interviews.

4.5.1 Analysis of Quiz Results

For the first time, automatic data collection was used, which increased the amount of quantitative data that could be analysed substantially. Below, the findings from each data analysis method are presented. There was one test group and one reference group, the difference being if they were allowed to consult the manuals or only use the feedback from the app.

Quiz Results and Pre-Test Data Analysed in Google Sheets

For results gathered and analysed within the Google Sheet, see figure 4.18. For a zoomed in version, see Appendix **??**. Early observations from the pre-test data when inserted into Google Sheets was that a surprising number of cells were left blank. One user had not done the pre-test (see column for coach 220), where some had left questions unanswered (most commonly "Do you own a company?" (should have used the word "business"), plus "Hours of preperation" and "Occations for a youth session" (there is a tendency this might be because they were not proud of their answers, because of correlations with low quiz results).

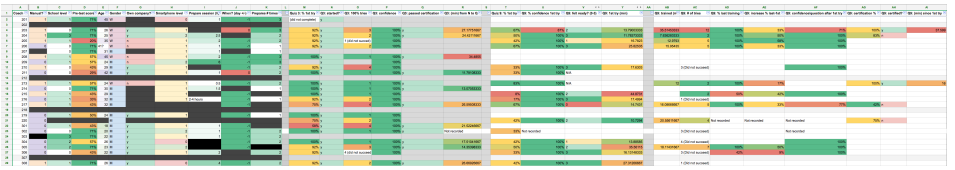

Figure 4.18: *The Google Sheet after merging the pre-test data and the summed quiz results data. See zoomed-in versions and explanations for each section below.*

	A	B	C	D	E	F	G	H	I	J	K
1	Coach	Manual?	School level	Pre-test score	Age	Gender	Own company?	Smartphone level	Prepare session (h)	When? (day +/-)	Prepares # times
2	201	1	1	71%	45	W		0	1	-1	3
3	202										
4	203	1	0	71%	26	W	y	1		0	3
5	204	1	3	71%	25	W	y	2	2.5		2
6	205	1	1	20%	35	W	n	1	2	-1	2
7	206	0	0	71%	41?	W	n	1	1	-1	1
8	207		0	71%	31	M		1	1		1
9	208	1	1	57%	45	W	n	1	2	-1	3
10	209	0	1	57%	24	M	n	2	6	-1	3
11	210	1	0	43%	29	M	y	1		-1	2
12	211	0	0	29%	42	M	y	1	1	0	2
13	212										
14	213	1	1	57%	24	W	n	1	0.5	-1	2
15	214	0	0	71%	30	M		1	1.5		1
16	215	1	1	43%	29	M		1		-1	
17	216	0	1	33%	32	M		1	2-4 hours	-1	1
18	217	1	1	43%	32	M				-1	
19	218										
20	219	0	1	50%	24	M	y	1		-1	1
21	220	0				M					
22	301	0	1	43%	18	M	y	1	1	-1	
23	302	0	0	71%	20	M	y	2	1	-1	2
24	303		3	71%	22	M	y	1	1	-1	2
25	304	0	2	57%	26	M	y	0		-1	2
26	305	0	2	71%	23	M	y	2		-1	
27	306	0	0	43%	22	M	y	1	2		2
28	307										
29	308	1	1	71%	26	M	y	0	4	-1	2

Figure 4.19: *Pre-test data*

	A	M	N	O	P	Q	R
1	Coach	Quiz 3: % 1st try	Q3: started?	Q3: 100% tries	Q3: confidence	Q3: passed certification	Q3: (min) from N to Q
2	201	(did not complete)	y				
3	202						
4	203	92%	y	3	100%	y	27.17751667
5	204	92%	y	2	100%	y	24.42111667
6	205	92%	y	1 (did not succed)	100%		
7	206	92%	y	2	100%		
8	207						
9	208	100%	y	1	100%	y	34.4855
10	209		y		100%		
11	210	92%	y	4	100%		
12	211	100%	y	1	100%	y	11.79108333
13	212						
14	213	100%	y	1	100%	y	
15	214	100%	y	1	100%	y	13.07053333
16	215	100%	y	1	100%		
17	216	92%	y	2	100%		
18	217	75%	y	4	100%	y	26.59508333
19	218						
20	219	100%	y	1	100%		
21	220	75%	y	2	100%		
22	301	58%	y	4	100%	y	21.52246667
23	302	100%	y	1	100%	y	Not recorded
24	303						
25	304	100%	y	1	100%	y	17.51041667
26	305	92%	y	3	100%	y	14.35398333
27	306	92%	y	4 (did not succeed	100%		
28	307						
29	308	92%	y	2	100%	y	26.66926667

Figure 4.20: *Quiz 3 answers.*

	A	T	U	V	Y
1	Coach	Quiz 9: % 1st try	Q9: % confidence 1st try	Q9: felt ready? (0-3)	Q9: 1st try (min)
2	201				
3	202				
4	203	67%	87%	2	13.79003333
5	204	50%	100%	3	11.78373333
6	205	42%	100%	1	16.7823
7	206	67%	100%	3	25.82535
8	207				
9	208				
10	209				
11	210	33%	100%	3	17.6303
12	211	33%	100%	N/A	
13	212				
14	213	83%	100%	N/A	
15	214				
16	215	8%	100%	2	44.8731
17	216	17%	100%	3	17.4994
18	217	67%	100%	0	14.7431
19	218				
20	219				
21	220	42%	100%	2	10.7294
22	301				
23	302	33%	Not recorded		
24	303				
25	304	42%	100%	1	13.88585
26	305	50%	100%	2	35.56115
27	306	33%	100%	3	16.13148333
28	307				
29	308	42%	100%	3	27.31266667

Figure 4.21: *Quiz 9 try 1 answers (correctness and recorded confidence ("Are you sure?": "Yes/No") on coach guide quiz 9 "Are you ready?", together with answering "How comfortable and ready do you feel right now to carry out session 9? (There is no right answer, just be honest with yourself!", alternative 1-4 ("Not ready at all", "Somehow ready", "Ready" or "Very ready and comfortable"). Finally in column Y, the time it took for the coach to complete quiz try 1 is given.*

	A	AB	AC	AD	AE	AF	AG	AH	AI
1	Coach	Q9: trained (m	Q9: # of tries	Q9: % last training	Q9: increase % last-1st	Q9: confidence/question after 1st try	Q9: certification %	Q9: certified?	Q9: (min) since 1st try
2	201								
3	202								
4	203	35.51453333	12	100%	33%	71%	100%	y	37.599
5	204	7.636283333	2	100%	50%	100%	83%	n	
6	205	12.9793	5	100%	58%	100%			
7	206	15.95435	5	100%	33%	100%			
8	207								
9	208								
10	209								
11	210		3 (Did not succeed)			100%			
12	211								
13	212								
14	213	12	2	100%	17%		100%	y	16
15	214								
16	215		2	50%	42%	100%			
17	216		1 (Did not succeed)						
18	217	16.08666667	5	100%	33%	77%	42%	n	
19	218								
20	219								
21	220	20.55611667	4	Not recorded	Not recorded	Not recorded	75%	n	
22	301								
23	302		3 (Did not succeed)			Not recorded			
24	303								
25	304		4 (Did not succeed)			100%			
26	305	18.11431667	7	100%	50%	100%			
27	306		3 (Did not succeed)	42%	9%	100%			
28	307								
29	308		1 (Did not succeed)						

Figure 4.22: *Quiz results on the training and certification of coach guide quiz 9 "Are you ready?". Column AB: how many minutes they spent in training. AC: how many times they pressed "Try again", retaking the wrong answers. AD: their score on the last training quiz they took. AE: the increase from their first training to their last. AF: recorded "Are you sure?": "Yes/No" after taking the quiz try 1. Finally, for the coaches that started the certification, their results are shown, together with the time the quiz took for those that got 100% correct on the first try.*

	A	AK	AL	AM	AN	AO	AP	AQ
1	Coach	Pre-quiz Q1	Pre-quiz Q2	Pre-quiz Q3	Pre-quiz Q4	Pre-quiz Q5	Pre-quiz Q6	Pre-quiz Q7
2	201	2	1	2	4	3	2	1
3	202							
4	203	2	1	2	1	1	2	1
5	204	2	1	2	1	2	2	3
6	205	2	1	2	5 (1,4)	1	2	1
7	206	2	2	2	3	1	2	1
8	207	2	1	2	1	1	2	1
9	208	4	1	2	3	4	2	3
10	209	2	1	3	1	1	2	1
11	210	2	2	2	1	1	2	2
12	211	4	4	2	1	4	4	1
13	212							
14	213	2	1	2	1	1	4	1
15	214	2	1	2	1	1	2	1
16	215	2	1	2	5	1	1	2
17	216	1	2	5 (1,2,3)	5 (1,3)	4 (1,2,3)	5 (1,2)	1
18	217	2	2	2	5	4	2	2
19	218							
20	219	2	1	5	5	1	5 (1,2)	1
21	220							
22	301	2	1	3	4	1	2	2
23	302	2	1	2	3	3	2	3
24	303	2	1	2	3	1	2	3
25	304	4	1	2	3	3	2	2
26	305	2	1	2	3	3	2	2
27	306	2	2	1	3	3	2	3
28	307							
29	308	2	1	2	3	1	2	3
30		85%	80%	89%	33%	5%	89%	55%

Figure 4.23: *Pre-quiz results per question and coach. The correct answers are given in green. The questions asked can be observed in Appendix B*

.

Missing cells was not as obvious with the app results, were users could not progress in a quiz without answering both the question and the confidence. However, none of the passed quiz 9 certification answers had been submitted. Thus, it was needed to add these from the manual recordings, which had been used as a backup in case anything like this would happen.

There were a number of quick insights that could be drawn before the parallel coordinates visualization, simply by looking at the data as a spreadsheet.

It was believed that smartphone users feeling like novices might have a disadvantage with the app, since they will not learn as fast as experienced users. The interactions shows however, between iteration 3 to 4, almost all of the coaches does not feel intermediate instead of beginners using the smartphone and the YoungDrive app. The quiz data verifies this, with no direct correlation between technical skills and quiz results (comparing column H "Smartphone level" 1-3) with column M and T (quiz scores on the first try for quiz 3 and 9). From the pre-test data, it can be seen that only CBTs said they didn't feel comfortable with smartphone (n=2, column H, of which there were 6 Youth Mentors and 14 CBTs). A reason might be age, as CBTs were older than the Youth Mentors. Also, youth mentors had higher school level than the CBTs. More probable, is that the experience of using a smartphone since iteration 3 has matured over time, and that they are now more confident.

From the Google Sheets quiz results data, it could be seen that there was a surprisingly low number of answers where the user answered the question without confidence (see column P, U and AF). This is not good for feedback purposes, as answers were actually often not correct.

First-hand insights before the parallel coordinates were that there was a strong correlation between pre-quiz results (column D) and quiz 9 try 1 (column T), and slightly visible also in quiz 3 try 1 (column M), but with more outliers. Also, with manuals there was a higher probability to finishing quiz 9 training and certification (see also findings from the parallel coordinates visualization below). The data shows small tendencies that being honest and deliberate during the training increases the likelihood that a coach can get 100% without faults, and has truly learned.

The final version of the app shows users can get 100% on quiz results much faster (column R, AB and AI) than the previous version, as the score board had been improved for iteration 4. Since the target group in Zambia and Uganda was different, it is hard to empirically prove if it went faster getting 100% with the possibility of repeating only the wrong questions, asking "Are you sure?", and providing individual feedback. Asking after the app does in iteration 4 does show however, that 100% of the coaches now thought the feedback was good for learning.

Bonus Result: Learnings from Observing All of the Quiz Results Also, from observing all of the quiz submissions (see sample in figure 7.1), more users had started a quiz without finishing it than anticipated. This speaks for usability issues (cancelling the quiz by mistake, like clicking the back button, or loss of internet access) or bad technical skills (needs to design to be more fail-safe) or lack of motivation (needs to design for unmotivated coaches as well. The need

group of "challenged coaches" might need other design than "ideal coaches".

Finally, a lot of users had done quizzes that were not Topic quiz 3 and Coach quiz 9, which might indicate high interest (if they did more than 2 quizzes) or confusion (if they did not do 3 or 9, but they did do other quizzes) during the app evaluation. This meant that on some aspects, there were less data than anticipated, (which was troublesome, as there were already few data points), and some aspects where there was more data than anticipated (on overlooked other quizzes).

Parallel Coordinates Visualization

Statistical analysis in R showed that none of the results were statistically significant to be notable using linear regression, or any of the other statistical methods detailed in the methodical framework. It was believed that statistical analysis could at least give insight into what to look after in the parallel coordinates visualization, but also for this, a larger sample would be needed. An interactive parallel coordinates visualization could give many more insights, more faster than the static presentations in Google Sheets or statistical analysis in R.

There is an almost infinite number of findings and observations that are interesting to look at via the parallel coordinates visualization. Even if there are no clear correlations (which are often found immediately when working with large data sets and parallel coordinates), tendencies can be found. As nothing is statistically significant (mostly because of the small data sample), the only thing that could be found are tendencies, and to analyze outliers. Then, these findings was critically analyzed to characteristics, and observations made from the app tests. As such, the data should be seen as indications of where future research is interested, and not as universal truths. For further research into the data, see figure 4.24 and the below sections.

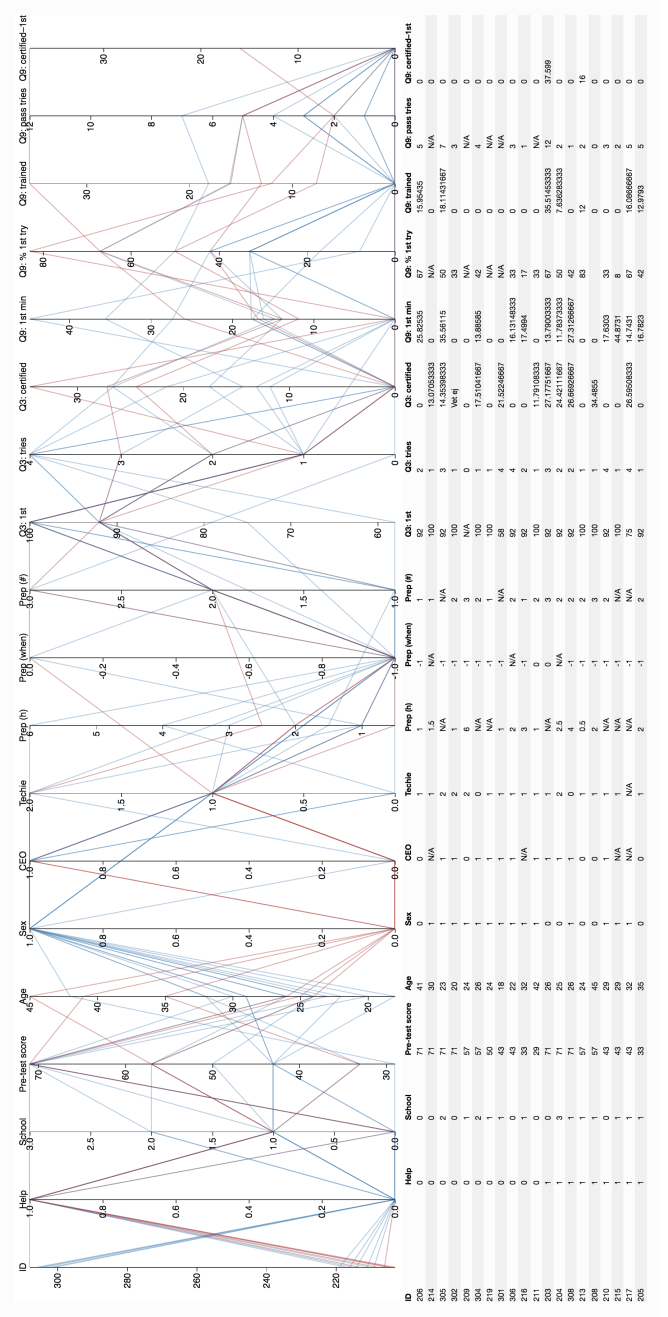

Figure 4.24: *The quiz results from iteration 4 shown in an interactive parallel coordinates visualization. The visualization is available on* `http://marcusnygren.github.io/youngdrive-parallel-coordinates/`.

Following the lines from column "ID" (for Coach ID), you can compare the difference in performance between Youth Mentors (301-308) and CBTs (201-220). Between quiz 3 and 9, there is no unison difference. In quiz 9 however, the Youth Mentors are top performers compared to CBTs, which goes in line with the project leaders opinion in the field that Youth Mentors are slightly better in the field than the CBTs. This could be explained also by that Youth Mentors only teach Young-Drive, while the CBTs also teaches other programs. It is important to note that there is nothing statistically significant to draw confident conclusions. Further research needs to be made, as a connection between how a coach performs in the field versus the app is valuable.

The results shows that the ideal coach, according to the quiz app, would be a woman, both in contrast and in unison with the literature mentioned in section 2.1. In figure 4.24 it can be seen that the female coaches (red lines) on average has better scores than the male coaches (blue lines), in spite of having less formal education. Further, she prepares more, is more aware of her own knowledge and has a better study technique, respecting the app feedback for meta-cognition and meta-memory. More than average, for example in quiz 9, the women have a higher lowest threshold, and a much higher record, than the men. two people were fast enough to get certified on the final quiz before the app evaluation ended, see figure 4.25. Both were women. Apart from gender, these coaches had higher quiz results, faster learning, and more honesty in "Are you sure?" than the others. Today, the balance between male and female coaches is reversed from what the data says: in Zambia only men have been hired, and in the data collection for Uganda, only 30% (6/14) were women.

Other social characteristics on the two coaches that passed the certification for coach quiz 9 is that were that both were CBTs (not youth mentors), and were in the middle of the age groups (24 and 26 years old). Regarding performance, they had a good pre-test score (57% or 71%), had top scores on quiz 3 try 1 and 9 try 1. Also, both of them used the manual, they looked at themselves as medium-skilled using a smartphone, and they prepared many times per youth session (2 or 3 times). In figure 4.25, a more detailed explanation is given. What didn't seem to matter for top performance, was number of tries for passing the training of quiz 9 (one coach did 2 tries on quiz 9, the other 12 tries), or time to pass training quiz 9 (35.5 minutes on the slowest versus 12 minutes on the fastest). Neither did it seem to matter when they prepared their session (1 did preparations the same day, 1 the day before). Regarding social characteristics, one had a business, one didn't, and their school level were both low.

Finally, it is interesting to observe the differences (and lack thereof) between the test group and reference group, by following the lines from the second column, "Help", of figure 4.24. In the test group ("Help" = 1), the paper manuals could be read before improving on the quiz results (not during the actual test) The the reference group ("Help" = 0) were only allowed to observe the right answers within the app, from the score board. In quiz 3, where almost all coaches had 92% or 100% immediately, there is no difference observable. However, in quiz 9, the hardest quiz, 5/7 that passed the training were in control group A, and 2/2 that passed the certification were in control group A. An explanation

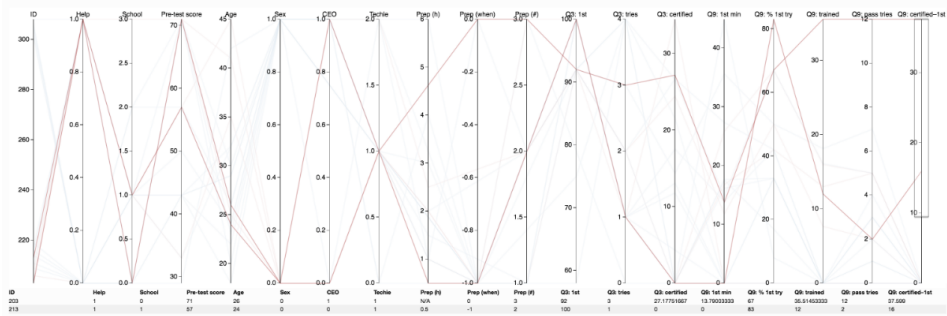

Figure 4.25: *The parallel coordinates visualization showing the characteristics of the two coaches that passed the quiz 9 certification. They were both CBTs, they could both could consult the manuals, they had a low education, while still having higher pre-test scores that the average. They were relatively young and were both women (in minority). One had a company and one did not. The one that inserted how she prepared for session, said she prepared the day before (versus doing it the same day), on three occations. Regarding quiz results, they both had a high score on try 1 of quiz 3 "Financial literacy" (1 or 0 errors), just like the majority, and it took them 1 or 3 tries to pass the training. For the coach where time is recorded for getting certified, it took her 27 minutes to pass the quiz (higher than the average). On quiz 9, it took 13 minutes to get 67%, the other coach (time not known) scoring 83%. While the coach with 83% passed the training in only 12 minutes and 2 tries, it was the coach taking 35 minutes to take 12 tries that could afterwards get all of the answers correct in the Certification.*

could be that the large amount of questions made the correct answers hard to memorize versus actually learning, or that the ones with manuals felt more supported or motivated because of the extra support. While these findings could be true for a larger sample, further research needs to be done. The same methods of data analysis are increasingly relevant with a larger data set, and there seems to be correlations and tendencies worth looking further into.

4.5.2 Analysis of Interviews

The answers given during the three group interviews has been summarized and clustered into three areas, see figure 4.26. Further, the answers within interaction design has been clustered into its four main principles, see figure 4.27. The questions regarding learning has tried to answer why a coach is correct or incorrect on a given question, see figure 4.28. Service design has been clustered into the two questions "When do you want to use the app?" and "When are you not able to use the app?", see figure 4.29. Zoomed-in versions of all of the areas are presented after the overviews, where the analysis of the quotes can be seen in its fullest. The quotes from these mind-maps are individual, which means that if another coach has had a similar thought, their quote is branched next to the following quote.

Figure 4.26: *The interview answers has been clustered into three areas: learning, interaction design and service design.*

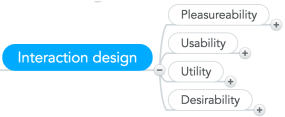

Figure 4.27: *The coach quotes regarding interaction design has been divided by four criteria: pleasurability, usability, utility and desirability.*

Figure 4.28: *The coach quotes regarding learning has been divided by if they give insight to why a coach has been correct or incorrect on a given question.*

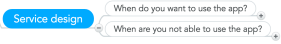

Figure 4.29: *The coach quotes regarding service design has been divided into when you want to use the YoungDrive app, and when you are hindered from doing so.*

Learning

By reading the quotes from the coaches carefully, it can be understood why answers can appear correct by the coaches (see figure 4.30), even if this would not be the case (see figure 4.31.

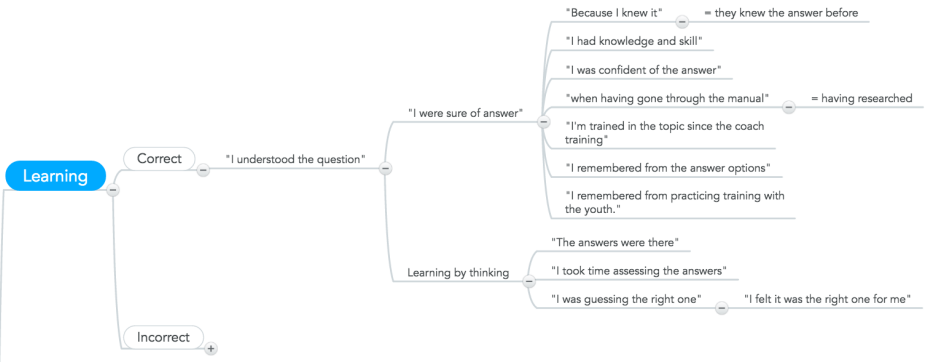

Figure 4.30: *Quotes explaining why a coach could give an correct answer on a given question. Either you were sure of the answers, or you made a qualified guess. A prerequisite for answering the question correct is that you understood the questions meaning.*

Figure 4.31: *Quotes explaining why a coach could give an incorrect answer on a given question. Either, the coach did not have sufficient knowledge to answer the question confidently (for a number of reasons, given in the figure), or the question was not understood correctly, or the app usability was a hindrance.*

Interaction Design

Most importantly, in the final app test, everyone thought the app was good and easy to use (n = 26). This is important, as this had not been the case in iteration 3. Regarding the four different areas of interaction design, positive remarks on utility are especially beneficial for learning. However, pleasurability, usability and desirability is a prerequisite for the app to be used by the coaches. For desirability, if the coaches are stimulated by using the app, two reviews were: "It felt good using the app" and "It motivated learning". The other three interaction principles (pleasurability, usability and utility) have been given their own mind-map below, see figure 4.32, figure 4.33 and 4.34.

Figure 4.32: *Pleasurability was 100%, as the app had a good design and well-motivated questions, however some coaches were frustrated by difficult questions in the app, or the time needed to complete a quiz during the workshop.*

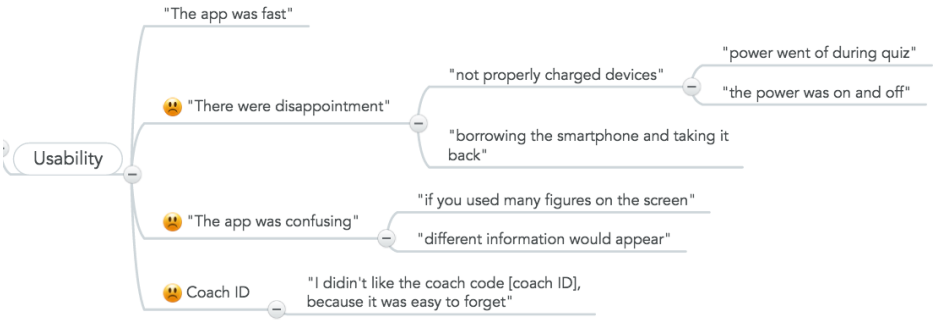

Figure 4.33: *Usability. As previosuly stated, all of the coaches (25/25) thought the app was easy to use when asked by raise of hands. However, in the interviews, some detailed comments regarding usability appeared. Regarding workshop issues, some of the devices' battery died during the workshop, or needed to be replaced. Regarding the app, some still mention too much information to appear at once, or that information are shown that the coach does not expect. Only one coach mentioned thinking that thee login was not user friendly, since the Coach ID was easy to forget. The Coach ID has since been documented by the local project leaders, so they can be contacted in such situations.*

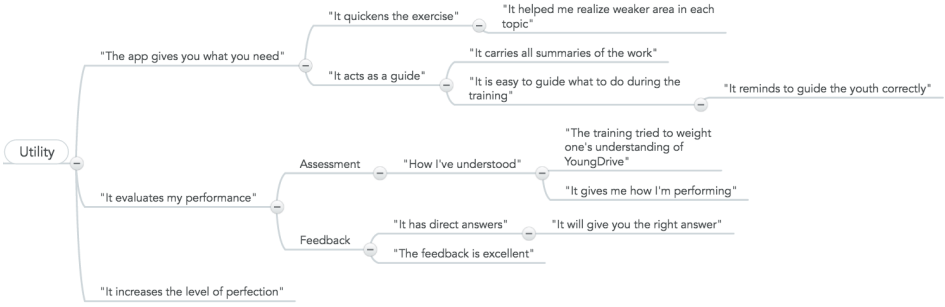

Figure 4.34: *Utility gives a valuable measure of what benefits the coach finds with using the app. Quotes regarding usefulness are: "It quickens the exercise", "It carries all summaries of the work", "It reminds to guide the youth correctly". For training, coaches conclude that "It evaluates my performance" and that "The feedback is excellent". One coach summarizes with "It increases the level of perfection".*

Service Design

It is important to understand if and when the app will be used by the coach, and if the environment of the coach in any way can hinder use of the app. For understanding the coach situation to these two criteria, see figure 4.35 and 4.35.

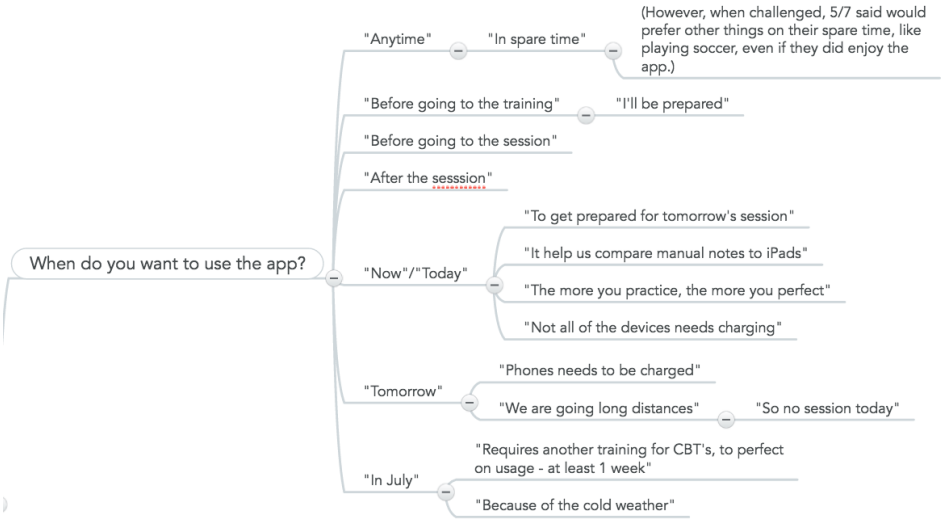

Figure 4.35: *There are very varying answers to the question "When do you want to use the app?". Most coaches would have liked to use the app immediately. Some of the coaches identified the need for devices, more training, or charging of devices in being able to do so.*

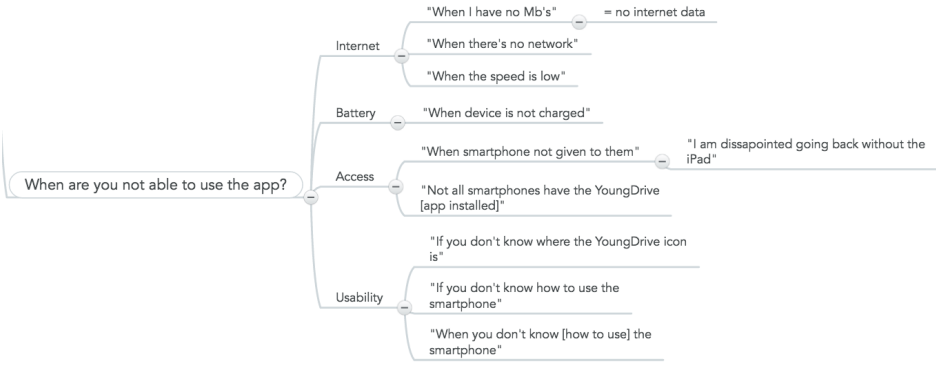

Figure 4.36: *Answers for the question "When are you not able to use the app?" are grouped into four segments: internet issues, battery issues, low or no access to smartphones, or if the app is not usable because it is not available on the smartphone, or the coach does not know how to use a smartphone.*

5

Discussion

The discussion section is framed by revisiting the five research questions. For each question, important aspects are considered, often comparing with the literature. This then leads to the conclusions of the master thesis, and future work.

5.1 How is the Development Affected by the Technical Possibilities?

As devices were limited, a goal was to make the app available on as many devices as possible. Creating a hybrid app using web technologies using Meteor made the app available both as a native app on Android and iOS devices, as well as on the web. On the other hand, this is not enough: the pre-evaluation showed that only 3 out of 16 had a smartphone today (Lönn, 2015c).

As internet is accessible but expensive and often used seldom, the app does not provide rich media or simulations, but focuses instead on creative design possibilities using multiple-choice, also having cognitive load and scaffolding in mind. The interviews and observations from iteration 4 details that the coaches are happy with the user friendliness of the app, and that the training in its current form has high value for the coaches. On the other hand, images could probably be used to lower misconceptions in language, and a future wish of the coaches is to have the manuals accessible via mobile as well, which includes both text and images.

Most of the coaches have been first-time smartphone users. Letting them continuously test and co-create the app has created a tailor-made app from their needs and conditions. It may be surprising how simple design solutions using text and clear visuals can provide rich learning feedback, mentioned by Nocol Nicol (2007). On the other hand, since it is so tailor-made, it needs to be exam-

ined if to make the app work in other countries than Uganda and Zambia, where design and technology preferences might be different.

That the app should work offline, and still be able to push quiz results when it gets online, has been a challenge. It can be hard to find good existing approaches for some technical platforms, but for Meteor plugins such as GroundDB proved very usable, since it is also automatic. In other apps in developing countries sometimes the user decides themselves when to push data, but in this case, quiz results are so small in size that it was unnecessary. This might be reconsidered in the future, for example if answers are no longer solely multiple-choice.

Below, reasons the development was negatively affected by the technical constrains are highlighted.

5.1.1 Online Data Collection was Needed Earlier

To test on all of the coaches in Uganda, it would have been preferable if data collection would have happened via the app instead of manually already in iteration 3, since there would be more than 10 test subjects, which had been the limit in Zambia. This was planned for, but technical implications with Meteor made it delayed. Done manually, not all data was recoded in iteration 3, which made it harder to draw conclusions from the quiz results. Both Lopez Lopez (2016) and Ropinski (2014) explains how visualization techniques (like parallell coordinates) are more suitable for large data sets. This can be read more about below.

5.1.2 Problems with Internet Access

In day 3 of the Zambia coach training in iteration 2, iOS no longer allowed uncertified app installs from computer: you needed to have paid a license even for unreleased apps, being a "Trusted developer". This stopped the app from being able to be installed on all the iOS devices, so that only the web version could be used. Thus, only the web app was tested from Wednesday and onwards. This was a problem, as the app regularly crashed at refresh because of low internet capacity. Sometimes, it was neeed to go to the other office where there was wifi, to refresh the webpage, and go back to the location. Going away from the training would of course not be viable Josefina, as she would miss valuable time teaching the youth. A similar situation appeared for the current author: while it was positive that these challenges were identified thanks to real-world testing, valuable time testing the actual functionality of the app was lost, with less feedback for continued development as a result.

5.1.3 Backwards Capability Issues

Upgrading from version 1.2 to 1.3 during Iteration 3 was a good example of technical limitations. It took a lot of time, but when it was discovered that version 1.2 did not work for old Android devices, the changes needed to be reverted. In

another project, new Andorid versions might have been acceptable, but here a "better" version of software was not viable.

Meteor 1.2 had several disadvantages: while it worked for all devices, it did not support React.js Meteor 1.3 was released, which promised a better developer experience, with JavaScript ES6 support, and access to Node Package Manager (npm), plus official support for React.js. In 1.2, only some npm packages had been adapted for Meteor, and tools such as Webpack could not be used.

The downsides was discovered after implementation: there were missing backward compatibility to the older of the Android devices. The backup would be the web version, but at the time of iteration 3, there were no Heroku build-pack for Meteor 1.3, making the website to crash. This was however fixed before iteration 4, which is why Meteor 1.3 was kept.

5.1.4 No Time Assigned for Writing Automatic Tests

The project would have benefited from passing automatic tests before doing user tests. While automatic tests were never written because of time constraints, since iteration 3, beta releases and production releases were separated into different domains, using Heroku's staging environment, with a different GitHub branch for each new iteration. Even so, doing automated tests would could have helped finding things that had worked previously but not in a new version, or finding bugs with new functionality like client-server communication. This would have made interactions with coaches more efficient, since the users would have been exposed to an app with less accidental errors.

5.1.5 Difficulties Comparing Quiz Results between Iterations

It would be highly interesting to compare the quiz results between different iterations of the app, to measure how much learning has increased. However, the educational range and knowledge between Zambia and Uganda is too large to draw such conclusions: while all of the Zambia coaches had 100% correct answers on quiz 3 "Financial literacy" (iteration 2), the same number for Uganda was 91.8% (iteration 4). This makes further conclusions very hard, more than from informed guesses and observations. See more about this in section 7.1.3. How to empirically measure and compare learning effectiveness between different countries could be interesting future work.

5.2 How is the Design Affected by the Contextual Constraints?

An insight is how quickly the coaches have increased their fluency with using the smartphone and the app. Even though the design at first needs to be very simple, as long as features are introduced slowly and as intuitive, the app could become more complex over time, to the point where no compromises needs to be made.

When it comes to design constraints in regards to overcoming cultural differences, learning from the expertise of local partners and technology companies can not be overestimated. They have been very willing to share previous mistakes, learnings and successes. This has saved a lot of time, and made the sparse amount of interactions and development time so much more efficient. Also, the value of getting to know the coaches on a first-hand basis has been greatly beneficial. The app has been designed together with them as co-creators, with a developed mutual interest and understanding, having a common goal of creating an app that works for their needs, in accordance with service design methodology (Stickdorn and Schneider, 2010).

The design was heavily influenced by starting the project with an iteration to truly understand the target group and context, and see the needs before starting thinking about ideas. It was good that this iteration did not have a digital focus, but even questioned if the best way to solve the needs would be an app altogether. Thanks to service design, everything in the app is informed by needs, which is an example of being a thoughtful designer (Löwgren and Stolterman, 2007) (Stickdorn and Schneider, 2010)).

Service design methodology allowed the design to be embraced by the contextual constraints. Since service design involves looking at the whole context, both the digital context and the physical context, it was possible to understand in what situation the app can be used, and the situation of the people using it.

When it comes to designing for entrepreneurship, the focus for this thesis was on the educators: both the teacher, and the YoungDrive coaches. This goes in line with the insights from Ruskovaara and Pihkala (2015), where the teacher seemed to be the main factor for a high-quality entrepreneurship education. However, contrary to Ruskovaara and Pihkala (2015), women did perform better than the male counterparts in the quiz results. Similar results have been found in other developing countries as well: from the pre-test, it is noticed that women puts more work into preparation, even if they have a lower school level. This could be one reason why their results where better.

More so, the interviews with coaches shows that the app has been fun to use, which was a recommendation from both YoungDrive and the research from Dickson and Weaver (2008) on recommendations for entrepreneurship education. Secondly, the app has focused on another best practice from Dickson and Weaver (2008), which is that the app should develop the competences as a mentor, enabler or coach. This can be noticable by interview answers such as "open's up ones mind" and that it "evaluates my performance" with "how I've understood", "how I'm performing" and that "the feedback is excellent".

From Computer Expert into a Digital Service Designer

The iterative study method supported the project in combining research and experiments, and doing constant improvements, to make increasingly well-informed decisions. For the current author, new-found skills has been acquired within:

- ethnology (getting to know and learn from people in a different culture)

- human-centered design
- design thinking
- service design thinking
- interaction design
- digital learning
- data analysis

It has placed high psychological pressure and leadership demands as a new designer, to:

- always be in charge of balancing all the different perspectives, with the end user's best in mind
- be able to change the planned process when new learnings or opportunities emerge (leading an agile design process)
- always implement new functionality from customer needs instead of designer or engineer bias
- continually design and run workshops and tests suitable for the target groups

The reason why this has been especially hard, is that simultaneously to learning design and technological skills, it has been in a different cultural setting than the designer is used to. This has also been extremely rewarding, at the same time exhausting.

Having Many People Involved in the Design Process to Adapt to the Contextual Constraints

The contextual constraints affected the design in the way that more expertise and guidance was needed than otherwise, from a diverse set of people. Service design research proposes to have a diverse team to build with a holistic perspective (Stickdorn and Schneider, 2010), and this has been followed.

In the project, working with service design in another culture has been one of the hardest and most exciting challenges. When the interactions for iteration 1 were cancelled the day before the trip because of local elections, experts could empower and affirm that this was not out of the ordinary. A learning was that it is only possible to plan to a certain extent, but then changing the plan in the last second, is needed more than when one's own culture. This support from experts has been very helpful, having several people familiar with working in a different cultural context before, which is one of the reasons Stickdorn and Schneider (2010) suggests making a project team multi-disciplinary.

To seek out people and situations that were not obvious, gave new insights into the work. The project partner Plan International, did not only back up the interactions and provide their own expertise, but also allowed interviews with experienced consultants from Grameen Foundation and Designers without Borders.

Another examples includes involving Expedition Mondial in the design process, testing the app on university students and refugee innovators and at startup hubs, and involving the local YoungDrive project leaders more than originally thought.

It was very valuable to combine having a diverse team with the designer having a clear direction and caring for the vision of the product. Otherwise, the product might have ended up in many different directions

The fact that the end users and stakeholders has been involved from the start, made them feel and have actual ownership of the product. This has many benefits, among others that *everyone* involved is satisfied with the *final* app, since their opinions and expertise has been taken into consideration and implemented. The fact that they can notice this further increases trust, and the the likelihood of them supporting future work, which goes in line with Stickdorn and Schneider (2010). To conclude, the design has been affected heavily by the contextual constraints, to the point where the end users are more likely to use the app as they have contributed to making a tailor-made product for themselves.

Spending Time in the Real-World Training

The original time plan stated that the interactions for Iteration #2 would have been in Tororo, and that it would not be possible to test the app during coach training whatsoever. However, during a Skype meeting with YoungDrive project leader Josefina, it was announced that it would be possible to participate in the coach training in Zambia during Iteration #3. A new work plan was created, which would allow travel to Zambia and to develop the app and participate in the YoungDrive coach training together with the coaches.

Now, it was shown already at Iteration #2, that if the app would have been created solely by the designer, it would have been assumed necessary with more functionality. Inventing the "field hackathon", designing and developing together with the users, meant that already for iteration 1, the purpose "Validate the coaches' level of knowledge during their education" could be fulfilled. The two following iterations could now focus on "Train the coaches on distance", and "Certify all staff". Also, without the 5 days of attending the training during the field hackathon, questions for each topic would not have needed to be created, and this would then have been a must-have for Future work.

The intense training in Zambia gave a lot of time to discuss and interact with the trainer, Josefina Lönn. One important contextual constraint that was noticed, was that Josefina did not want to be replaced, but appreciated having the YoungDrive to the point where the app should not replace her, even if it in the future would benefit YoungDrive in terms of for example monetary reasons. This might not only be personal preference, as Ruskovaara and Pihkala (2015) claims how important the role of the teacher is for effective entrepreneurship education.

Without visiting the Zambia training, it would have been much harder to focus on the coach training purpose of the app, since the Uganda coach training was already over. A consequence might have been that the app would have been more focused solely on "Train the coaches on distance". This shows that spending time in the real-world context of the situation you are developing for, is very

important. - especially when it is unfamiliar to you. This goes in line with the service design principles (Stickdorn and Schneider, 2010).

5.2.1 Having Low Scores on the Quizzes Might Lower Motivation

Interestingly from iteration 4, the coaches scoring from 0-53% on quizzes included two coaches being top performers in Iteration 4, but also the two coaches that did not show up for Iteration 4. Further research could be done, if the difference was on motivation or fixed versus growth mindset. It could be for some, that low quiz results in iteration 3 led to not wanting to use the app any more. The app was designed to be more empowering in iteration 4, but naturally, the app could not be tested on the two coaches that did not show up.

5.3 How Can Test Questions be Developed to Support Entrepreneurship Learning?

The problem identified with multiple-choice questions, regardless if the recommendations by Nicol Nicol (2007) were taken, is that they first could only measure lower-order learning objectives, see figure 2.1. While entrepreneurial *knowledge* objectives might be considered A-B 1-2, building entrepreneurial *skills* is definitely related to C-D 3-6.

When assessing the first question sets according to Bloom's Revised Taxonomy (Krathwohl, 2002), some characteristics were shown, which guided future creation of questions. Most notably, to reach C-D 3-6 on Bloom, there were some techniques: intelligent multiple-answers could encourage the coach to *evaluate* instead of using process of elimination or encouraging guessing. Putting the coach in a coach scenario (how to act in X situation?), the coach could be tested on a *procedural* and *metacognitive* level to *apply*, *analyse* and *evaluate* skills, and get feedback. The average increase by topic from topic from iteration 2 can be explained by that the teacher formulated new question sets after each day of training, getting feedback from the coaches and gradually understanding how to formulate questions to assess a certain learning objective. She was also increasingly informed and encouraged to assess questions according to Bloom's Revised Taxonomy throughout the week. When constructing the new quiz for iteration 3 and 4, "Are you ready?" for week 9, the teacher had a specific goal given to her to score higher on Bloom's Revised Taxonomy. This was also the quiz where coaches had the most difficulty getting a high score.

Previous research by for example Nicol (2007) had already shown how multiple-choice can be powerful, for example by following the principles in figure 2.3. The same articles mentions the approach taken with using a confidence-meter similar to "Are you sure?". These recommendations have been utilized.

Some bad questions have still existed, where coaches did not understand and failed to interpret the question, because too advanced English language was used. By receiving coach feedback and analyzing quiz results in-between iterations,

many improvements have been made regarding formulations. This is in accordance with agile methodology and SCRUM, where the value of testing the product with the intended target group iteratively is acknowledged (Kniberg, 2005).

Regarding testing and improving the quality of questions, the initial plan was that YoungDrive would only produce questions for two YoungDrive training weeks, not all 10. To have questions for all of the weeks have greatly benefited the master thesis, and increased the value of the final product. If not all quizzes would have been developed and tested, this would have been a Future Work.

From a question assessment, it is shown that all of Bloom's levels can now be reached via the app, but *create*. This is because users can not create anything in the app. Secondly, since the multiple answers are shown immediately to the user, they are not encouraged to apply their own thinking to the question, before seeing the alternatives.

5.3.1 Constructing Good Questions in Entrepreneurship

Consider the entrepreneurship topic question "What is financial literacy?" to Bloom's Revised Taxonomy (*conceptual* and *remember*) (Krathwohl, 2002). A learning is that to simulate a procedural environment for the coach, a question can effectively be presented as a scenario: "It turns out that 10 youth have not carried out the business action, what should you do?" (*metacognitive* and *evaluating*).

However, there are several traps that the person formulating the question and answer alternatives can fall into in the case of multiple-choice, where a good question might be de-amplified because of the answer alternatives (Nicol, 2007). Consider the coach being asked to give business advice to a fictional youth named Adam: "Adam wants to start a business that is based on a product. which business should he start?". In this case the coach has before been given questions on what a service and product is (factual remember), what the difference is (factual understand), and been given examples (conceptual analyze). Now, the skills are being put to a procedural test. If the answer alternatives are obvious (or memorized), the learning will be lower than scoring high on Bloom's Revised Taxonomy.

To construct high-quality answer alternatives, all of the answers must be evaluated and considered. In such cases, multiple-choice versus open-ended questions can actually amplify learning, by reaching *evaluate* on Bloom's Revised Taxonomy or by strengthening *learning by thinking* (section 2.3.4). In the case of the previous question valid alternatives for the coach to consider might have been: "Start a salon", "Start selling soap", "Start a bricklaying business". The coach must evaluate if each alternative is either a service or a product.

5.3.2 Necessary Improvements to the Multiple-Choice Design

It is still hard to score high on the knowledge and cognitive dimension using techniques such as multiple-choice with entrepreneurship and coaching. This is however necessary, if the app should reach the learning objectives of YoungDrive. This demanded additions to the multiple-choice design, and not solely content. Such design ideas was "Are you sure?" and giving individual feedback, both of

which encouraged metacognitive thinking. Ideas for future work can be read about in see section 7.4.

5.3.3 Learning Effect from the Questions

To the largest extent, the questions have been praised, in regards to formulation and challenge, which can be seen in figure 4.28 and 4.27. The lack of a post-test makes it hard to see if the test questions in the app has a real-world effect. Ideally, the post-test in Uganda would have been to observe coaches having their youth session, and compare their correctness and confidence behaviour when not having used the app.

The lack of data (not least from the pre-quiz), makes it hard to draw reliable conclusions from the data. Regarding the pre-quiz paper submissions, the submissions should have been checked for blanks before handed in. Another mistake was that school level was not always exactly specified, and this means that school level and quiz results might have a stronger correlation that can be shown now.

Regarding recording test question results, this should have been done manually already in iteration 3. The fact that it was done automatically in the app for iteration 4, made so that more data could be recorded, more reliably, than when the project leaders filled them in by hand whenever a coach raised her hand to say she was finished with a quiz. A learning effect is indicated partly by that in some cases coaches who had a low score on their first try with a quiz, after the training could pass the certification test, getting 100% in 1 try. This pattern needs to be more closely examined with a bigger number of testers.

5.4 How Does Design Affect Usability and Learning done via the App?

The quiz design is made with Kathy Sierra's model of deliberate practice in mind. Sierra (2015) Depending on the combination of correct and sure, the coach knowledge for each question is put into three different buckets: "Can't do", "Can do with effort" och "Can do effortlessly". By retaking the "Can't do" questions with Try again, reflecting on the "Can do with effort" questions (for example: correct but unsure), and waiting to test the "Can do effortlessly" questions until the certification, all questions are eventually put into "Can do effortlessly".

The coach can choose to leave the training without doing the certification quiz, later repeating the test. This is to make the learning self-directed and just-in-time, and to allow the coach to do it's own scaffolding. If the coach uses the app before a youth session, and has a low result, the coach will get feedback as such (for example: "Nice effort, but you still need to practice and prepare yourself even more! How can you do that?"). The goal is to allow the coach to move knowledge from "Can't do" to "Can do effortlessly" to "Certified" in a pace that suits the coach. This might be neccessary, instead of a one-fits-all solution, as the coaches' preferences are so different.

The learning goes faster in iteration 4 than in iteration 3, which is most notably shown by the fact that the speed from try 1 into getting 100% in 1 try has increased. This is largely because usability issues has been addressed, and because design choices has been made that stimulates and makes learning more efficient. Most notably, the score board has been improved to show which questions the coach is correct and sure of ("Can do effortlessly"), unsure but correct on ("Can do with effort"), and were wrong on ("Can't do"). One mistake made, was not to think about reducing cognitive load (Sweller and Kalyuga, 2011) until iteration 4, when usability issues were identified as serious problems in iteration 3. For the next iteration, the intrinsic load was lower due to more helpful feedback, the extraneous load was decreased by making the expected behaviour in the app more obvious (like "Try again" after having finished a quiz with errors). Finally, germane load was increased by the feedback of "getting certified" being related to getting ready for having a youth session, which was more relatable to the coach than only receiving a score and a medal (Sierra, 2015).

The YoungDrive app is the first known application which uses a confidence metric (in this case "Are you sure?") for the student's own sake, and not only assessment, like in the case study detailed by Nicol (2007). The effect is reaching meta-cognitive on Bloom's.

5.4.1 Benefits with Confidence Level and Correctness in Combination

If a coach is wrong and sure about a lot of questions, it might be the indication that the coach is teaching the wrong information to the youth, which might potentially hurt hundreds of their youth's businesses. If the coach is correct but not confident, it could be considered a guess, which is strengthened by the interviews in iteration 4, see section 4.5.2. In Sierra's framework of building expertise, it would be called "Can do with effort". In the app during iteration 3, dampened for iteration 4, there were troubles when coaches passing the training with too many correct guesses, knowledge that were not yet "Can do effortlessly". This meant that they would fail the Certification, because they could not answer reliably, providing the wrong alternative. This in turn meant, that they were put back into training instead of getting a medal, which was not motivating. The conclusion was that the coach should not start the Certification mode before being truly ready.

One solution could have been that "Improve" would not only include repeating wrong answers, but also answers where the coach had been correct but said "No" on "Are you sure?". However, coaches seldom believed they would be wrong, or at least did not determine it worthwhile to be honest. Some preventative measures were taken to try to make the coach more honest. Unfortunately, for iteration 3, only some coaches took notice of number of tries as an indication that they should pay more attention. For iteration 4, the design had improved by giving coaches minus points for being sure and wrong, but here as well, not everyone paid attention. This goes in-line with research both from deliberate practice being a desirable difficulty (learning being hard might make the coach more likely

to want to "cheat", or take an easier route like guessing instead of putting more effort).

The bigger problem was that the knowledge were not yet reliant, not that coaches were not honest. The solution for iteration 4 was to improve learning in the app, partly by a more personal score board with feedback, which could show the coach which kind of questions she needed to repeat ("Can't do" or "Can do with effort"), but instead of forcing the coach to redo correct guesses ("Can do with effort"), she could reinforce the correct answers by personal feedback. Sierra Sierra (2015) calls this "high pay-off tips", which can be very effective.

5.4.2 Deciding on Learning Methods

There was a lot of work behind choosing the learning design methods in Iteration #3. The way to progress, was to brainstorm various solution, discuss them with experts, and then create trigger material and test some of these approaches.

Retaking questions that were wrong ("Try again", called "Improve" in iteration 3) was inspired by deliberate practice Sierra (2015), and is already common in e-learning driver license software to learn traffic signs or how to act in various situations.

Showing the coach how many quiz tries they have done, was inspired by Linköping University's work with the e-learning tool NTA Digital, where they reward students with badges for getting 100% in few tries. Their goal with this kind of "gamification", is to reward students for studying before taking the test. Similarly, for the training mode, the coach seeing number of tries was a method of studying the correct answers more thoroughly. For the certification, where the coach was supposed to have trained before taking the test, badges worked with the same purpose as NTA Digital, to reward students that had studied properly.

"Are you sure?" was inspired by a Swedish teacher, and has been used before by others (Nicol, 2007). It has then been used to determine if a right answer should award a point or not. Nicol (2007) and the digital pedagogy advisor for this thesis, Henrik Marklund, suggested that the teacher had overlooked a learning benefit of this approach: the student reflecting on their own knowledge, which is proved great for learning. This was extended in this thesis where personalised feedback has the goal of the coach getting both confident and correct. In a school situation, this might not be necessary, but in the YoungDrive context, the purpose of the app was to build both correctness and confidence with the material. To pass the certification after training, getting 100% without faults, made the coach feel that confidence, at the same time reassuring the teacher that the coach had learned the material.

After iteration 4 a bonus test was made with Plan Tororo staff, which showed the relevancy of the certification mode: one group that were 100% correct on the first try, did get 100% correct on their second try, meaning guesses had been present. On the other hand, a person have 1 wrong answer, passing the training on the second try, did then pass the certification on the first try, with confidence. It can therefore be determined, that when all of the answers are answered correctly, after having gotten all answers correct once, that the coach has both cor-

rect information and confident - this is a good example of deliberate practice: the information has gotten reliable.

Other learning design methods were considered, as previously discussed. In a low-fidelity prototype in iteration 3 it was assessed how multiple-choice answers compared with using flashcards in a think-aloud test, see section 3.4.5. Challenges including flashcards in the high-fidelity prototype were that the coaches had no previous knowledge of typing on a keyboard, and analysing recorded answers would be too technically demanding. The integration and benefits of flashcards was however discussed in various ways, see chapter 7 Future work.

5.5 How Can Users' Feedback be Used to Inform Modifications of the App?

The number one time waster, has been spending time building ideas that is not informed by the users' feedback, and not immediately realizing the mistake. Such ideas will not be used, are often not supported by research, or does not fit the context. This is the same finding that Stickdorn and Schneider (2010) has made, and is a violation from service design principles of making the process user-centered. Luckily, almost all of the work with the master thesis has indeed been user-centered.

Getting enough feedback to evaluate each component of the app has been important. Enough feedback, means that modifications to the app can be made with confidence. The faster this can happen, the faster the app gets better. This is why the field hackathon, and service mini-sprints, were so beneficial, see 2.7. It ensured that fast iterations could happen with enough feedback to provide effective modifications and additions. The use of service mini-sprints was so effective that the interactions in Tororo could be expanded with even more days, as long as this is properly planned for.

As long as the users understand the purpose of the app, and how it should benefit them, they can be part of the creation process, not only the testing. This is an extraordinary opportunity, that should not be underestimated, as shown in case studies in **??**, although these projects has not had a digital focus. It is found that deeply involving the users and getting to know them is essential to understand the true needs of the coaches, and can greatly inform modifications of the app.

The app evaluation shows that not only should users' feedback inform modifications of the app, they should be the *basis* of the app. To do this, this report has provided but one approach, via Digital Service Design. It is also a service design principle, making the process co-creative Stickdorn and Schneider (2010).

The presence of different perspectives (from users, stakeholders, experts, and the designers care for the vision) lead to a holistic view when designing, inline with Stickdorn and Schneider (2010). The designer can balance these, as long as the views are present during the whole development process, and not only in the end of the project. For the questionnaire made for iteration 1, stakeholder views were very important. From their views, the questions could be improved and put

into context, and lessons were learned which otherwise would not have happened. In iteration 1 it was felt that maybe too much time was spent with stakeholders, almost to the point where the process where no longer user-centered.

By asking Why-why-why continually, the true needs of the coaches could be gradually exposed. Similarly, by hosting co-creation workshops, the coaches can be part of designing lo-fi app prototypes that addresses these needs. The role of the designer then, becomes putting these ideas into a context, and comparing with best practices and relevant research (in this case, most notably learning design), modifying, developing, and testing. As Löwgren and Stolterman (2007) says, the output will not be better than the designer carrying out the process. To embrace the role of a socio-technical expert, more so than the computer expert or political agent, has been one of the enablers for co-creation. Service design thinking and methods, gave a framework to have all of these perspectives in balance and consideration, always with the end user as the most important person.

5.6 Summary

As of now, the app can not be said to evaluate how good the coaches are with teaching. Neither, it helps them assess their ability to learn (it sure would be valuable for coaches to track their ability to get better and remember, see chapter 7 Future Work). The app does however help the coaches to assess their knowledge level and what areas they need to train on.

Moreover, the coaches seems to like the app, which is important. There are tendencies that the app works better for some coaches, especially those who take time to reflect on the feedback given by the app. This speaks for designing the app for different need groups. If motivation and confidence is high, it could indicate them becoming better teachers, but as of know there are no evidence of this more than that the coaches themselves has identified that confidence is important for having a good youth lecture.

A bonus result is that the app allows also the teacher to assess the knowledge level of the coaches on a day to day basis during the training, giving insight into how well their teaching has been received. The teacher can use this data to understand what they need to repeat the following day, or if adapting their teaching will lead to higher results. Allowing the teacher to analyse their test questions according to Bloom's Revised Taxonomy created awareness of what knowledge level they were assessing the coaches with by their questions, and motivated matching the formulation of a question to the knowledge and cognitive process dimension suitable for the educational objective.

6

Conclusions

In response to the research questions, the master thesis has:

1. Contributed to the domain of entrepreneurship education in a developing country context
2. Demonstrated how certain technical constraints and design constraints can be overcome in a developing world context
3. Provided methods of investigating usability and learnings with a digital training tool in the real-world entrepreneurship training context
4. Created new methods in service design, when co-designing digital artefacts in a developing country context

Below, these four items are explained in greater detail, tying the report to a close before suggesting future work in the final chapter.

6.1 Contribution to the Domain of Entrepreneurship Education in a Developing Country Context

This research shows that an app can be effectively used during and after training to assess and learn entrepreneurship. Furthermore, the app can also prepare coaches before youth sessions, training them both in entrepreneurship and preparing their lessons.

In *addition*, this has been done in a developing country context, with coaches having no prior smartphone experience. The app is used for coaches in conjunction with a physical entrepreneurship training today, and does not yet address the youth themselves, or replaces the physical training.

However, the research shows that both the teacher and the coaches greatly appreciates the support the app has given them. As entrepreneurship to its nature is practical in many aspects, one finding is how multiple-choice questions can increasingly simulate real situations by meta-cognitive feedback, using design solutions such as asking "Are you sure?", giving personal feedback, and formulating questions as scenarios.

6.2 Demonstration of how Certain Technical Constraints and Design Constraints can be Overcome in a Developing World Context

In this project, it was decided to design the app *together* with first-time smartphone users. Co-designing the app with the coaches and learning their perspective on overcoming the technical constraints was a valuable methodology.

Apps developed specifically for rural areas has previously been built that has overcome rural areas' challenges with access to smartphones, internet and electricity. This learning app is designed to consume low battery and internet by not being media-rich but on relying on text. It is also internet-aware by storing app data locally until internet is turned on. Until every coach is equipped with a smartphone, the app has been deliberately designed to work as a compliment to the physical coach training. After the training, it is not mandatory for a coach to use the app to help prepare for a session, even if it would improve youth session quality.

The workshops shows that coaches are willing and enable to pay for devices by micro-loan or utilizing the savings group, if granted access. Today, they have no such means by YoungDrive or the project partner. YoungDrive can cope with using the donated devices during the coach training, but needs to determine on a strategy to enable smartphone access to the coaches. A conclusion is how in many cases creative design can overcome technical constraints, but not always.

6.3 Provided Methods of Investigating Usability and Learnings with a Digital Training Tool in the Real-World Entrepreneurship Training Context

To investigate usability, observations using think-aloud in the real-world entrepreneurship training context proved the most effective. In big groups, data could give tendencies on common problems that needed to be addressed. In smaller groups, ideas and precise feedback was more common. It helped having a framework to compare usability against, in this case the interaction design principles of desirability, utility, usability and pleasurability (Clatworthy, 2010).

To investigate learning in the real-world entrepreneurship training context, literature research and data analysis of quiz results was highly beneficial, but

mostly when put into the context of the observations made in the real-world entrepreneurship training context.

Surprisingly often, comparing research and expert opinions with what coaches thought was best for learning, was in unison. When this happens, the self-confidence of the designer can increase, and the designer can be more daring and experimental.

Having much testing and co-creation was a very rewarding approach, as is discussed in section 5.5. To make it truly testable, lo-fi and hi-fi prototypes should be used instead of hypothetical questions. While research before starting to develop is a great start, it was trying different solutions, all based on user advice, expert opinion and research, that gave great results.

6.4 Created New Methods in Service Design, when Co-Designing Digital Artefacts in a Developing Country Context

Short iterations is a challenge, especially when time is sparse and the culture is different than one's own. Often in projects, testing is overlooked. In this project instead, service design helped to look at the users as not only testers, but as co-creators. This is not new in itself, but the creation of Digital Service Design, and methods within this discipline, were new. Examples include service mini-sprints and field hackathons.

Interactions with the coaches always included Service Mini-Sprints (see 2.7.5), allowing the most important feedback and insights to be addressed in the app and workshop formats already the next day. This allowed for very effective use of the sparse interactions with the coaches, and there was no need to wait for the next iteration to test the coaches' feedback.

The "field hackathon", in which the app was refined and tested each day of the coach training, was the most rewarding research and development opportunity of the whole master thesis. More than the opportunity to develop the app with the coaches, it had the extra benefit of giving the opportunity to observe the YoungDrive training and understand the coaches.

To use a service design approach when co-creating digital artefacts in a developing country context proved highly effective, and it is recommended that this area is further studied. It does demand bravery of the designer to get to know the users well (Löwgren and Stolterman, 2007), and design for their needs and dare to question the client as Stickdorn and Schneider (2010) encourages. In the end, as this project and Stickdorn and Schneider (2010) suggests, including both the users and the client will often make them more satisfied with the result, feeling both increased ownership and that the product feels and indeed has been designed from their needs.

7

Future Work

The future work section is divided by research question, and proposes additions that would strengthen the app goals (see section 1.1).

7.1 How is the Development Affected by the Technical Possibilities?

Some of the wished development were too complex to be implemented during the master thesis. Below, it is described what future work is wished in the area.

7.1.1 Data Analysis Improvements

Today, data analysis of quiz results takes a tremendous amount of time, and new results are thus not instantly accessible to the stakeholders. The data needs to be acquired into Google Sheets, and then needs to be enhanced in several ways to be filterable and visualized, see section 3.5. The raw data when inserted into Google Sheets can be seen in figure 7.1.

A lot of future development time should be spent so that most of this work is made automatically. Some of this work, is related to the way that the data is saved. Today, whether the coach was correct or incorrect on a question, and how many coaches answered that question incorrect after the first try, needs to be done by process of elimination, because the data is not properly structured in the database.

There were also a small number of errors with quiz submissions in iteration 4. Most notably, certification tests for coach quiz 9 were not submitted, which is why the paper submissions proved very valuable as a backup. To discover more errors regarding offline-online functionality, is important, as it is cumbersome and time-costly to test these manually. A good way to discover such errors, would

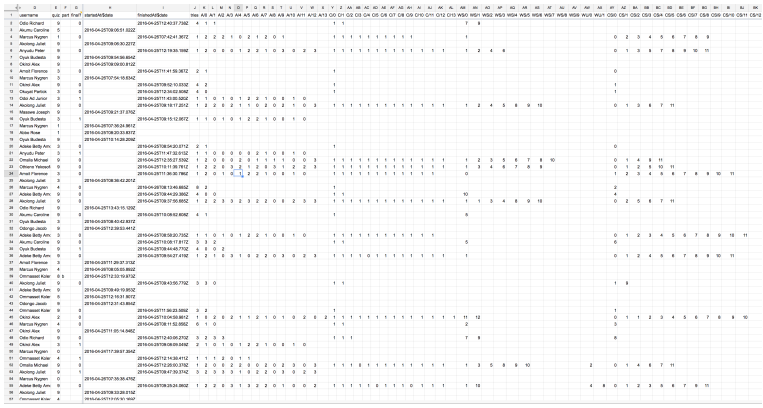

Figure 7.1: *The raw data of the quiz results when inserted into Google Sheets. The purpose of the image is to explain how much work that is needed in order to process the data into insights.*

be automatic tests (or regression tests), so that the app can be used by the coaches with confidence, without extra personnel present, checking that the app works.

7.1.2 Code Quality

Since the app will be continually used and developed by others in the future, code quality is important. React.js makes it easy to structure the code in a way that gives a new developer a good overview of the different components, and its functionality. Even so, refactoring the code into smaller components would be a good idea.

To increase speed of the app, refactoring the code to ensure that loading of assets happen more effectively (especially quiz questions), and also that data is cached in a logical way (saving necessary information so that it does not need to load the same information again, and vice versa), would be helpful. As the YoungDrive program grows and more people will use the app, there will be a lower tolerance for the app being slow. The app, especially on native, takes a long time to load, mostly because of asset loading.

7.1.3 Availability in More Countries

During the Zambia tests, the coaches were given a Zambia version of the app, and in Uganda, they were given a Uganda version. In the future, it would be advisable if the coach ID used for login would affect if the Zambia version or the Uganda version should be shown.

As long as the users of the app, such as the coaches in Uganda and Zambia, does have sufficient English skills, it is acceptable that the app is solely in one language. As more countries are introduced, the app should be available in different languages.

A challenge today is that the Zambia uses a newer manual, and e.g. the national currency is different than Uganda's. That Zambia has updated manuals, means that some questions are new, some are removed, and some are reformulated. But the teacher still wants to be able to compare quiz results between the different countries. Today, the quizzes were simply replaced between versions. This has the advantage of being easy to implement, but the clear disadvantage that the teacher can not compare quiz results between countries.

A good solution to internationalization of quizzes would be that each question in the database should include an unique ID, with different texts depending on country and coach manual. This would allow the teacher to keep track of some coaches have difficulties with a question, regardless of location.

7.1.4 Availability on All Android Devices

Since the introduction of Meteor 1.3, older Android devices are no longer supported. For the YoungDrive coaches, this is an issue, as most often devices use older Android operating systems, and does not have great performance. Today, a Meteor 1.2 version of the app is available on the AppStore. Also, the newest version is available on the web. However, in the future research should be put into if the app should abandon Meteor 1.3 altogether, or if there is another way to ensure backwards capability.

7.1.5 Availability on Feature Phones

As far from all coaches does not currently possess smartphones, few coaches will currently be able to use the developed app for preparing their youth sessions. Therefore, what was discovered in iteration 1, that all coaches possess feature phones, could guide the development of a SMS-based service. While this was not viable for the scope of the master thesis, research has been made into related work. A recommendation is to try VOTO Mobile (Mobile, 2012), which supports multiple-choice questions and internationalization. Today, this solution has been used mostly for doing evaluations in rural areas, via automatic phone calls where the caller can be given responses in text. However, research on using such a solution for educational purposes seems promising.

7.1.6 Reaching Statistical Significance

There are multiple examples how data analysis on a larger data sample could benefit. More than YoungDrive being able to take more informed choices regarding the future development of the app, it would strengthen answering the research questions. One such example is "Is the coach learning?". With enough data, it would be interesting to analyse the impact of feedback on getting the correct answer the next time. For example, you could compare if getting feedback on being wrong and unsure leads to better results than getting feedback on being wrong and unsure.

7.2 How is the Design Affected by the Contextual Constraints?

There were many things to consider when designing for the contextual constraints. Below, aspects that should be taken into consideration are presented.

7.2.1 Replacing the Teacher

In iteration 2, Josefina (the teacher) mentions that she does not want to be replaced by the app. However, there would be many benefits to YoungDrive if the coach training could happen 100% digitally. When the app was tested with refugee innovators in iteration 2, several of them asked if it was possible for them to use the app exclusively to train themselves in entrepreneurship or starting a YoungDrive group.

How could it be done in practise? Also, when the teacher does not want to be replaced? In practice, a freemium model could be proposed, where it is possible to take the coach training for free digitally, but pay for a physical training. Currently, this contextual constraint has affected the app in the way that it should complement the physical training, and ease the burden for the teacher.

7.2.2 Scaffolding the Coach Guides

Josefina says after iteration 2 that it is indeed valuable if the training prepares the youth more actively for holding youth sessions, an insight that was discovered in iteration 1. She mentions two challenges with introducing coach guides for each topic. In the end, a solution is suggested:

- The coaches are not ready the first day, as they have not gotten used to the app yet. *As such, they should be introduced in the middle of the training.*
- A recurring issue is that the Friday, the last day of training, should be dedicated to preparing a session, but the time has never been there. If so, the coach guides will not be used. *One idea could be to make the topic quizzes smaller, and mix topic questions with coach guide questions.*

7.2.3 More Time Designing the App for Different Need Groups

Already since iteration 1, different need groups have been identified. It is shown from the tests that the idealistic and realistic coach might be more probable to have a growth mindset, where challenged coaches might have a more fixed mindset. Continuing to use growth mindset feedback in the app is crucial according to Dweck (2014), who found her method made lower and medium achievers also played until the end, while the lack of such feedback only kept high achievers.

Also, there are tendencies that different need groups are more present in different countries. While a lot of research was done about the cultural dimensions of Uganda, the research done on the Zambia and local Kabwe population was more sparse. It was known that the socio-economic differences were large, but

not much more. For future work, it is recommended to put more work into how the local and national culture in a country affects the mentality of the coaches and the design. It is not possible to assume that the Zambia and Uganda culture should be similar, and it will be similar with other countries in other developing countries, or elsewhere.

Since doing app development together with the coaches has been so beneficial to discovering different needs, a wish is to have done more so with the coaches in Tororo. While in Zambia, the development was done in the real-world training, which was superb. In Uganda, more of the trigger material could have been created in Tororo instead of Kampala. Even if it would have been more costly and internet is slower, it would have been valuable being closer to the coaches.

7.2.4 Training the Coaches with Using a Smartphone

An additional insight from the smartphone test in iteration 1 was that using a smartphone operation system like iOS or Android needs to be made as easy as possible. This to avoid confusion with things like not finding the YoungDrive icon, or accidentally hitting a button, or click the power button: all of which relates to ease of using the operating system, and not to the app in itself. A lot of training is needed to avoid errors, and should be taken into consideration from a service design standpoint.

In iteration 3, such a co-creation workshop was held after the app test. This resulted in that for iteration 4, all of the devices had the YoungDrive icon on the home screen of the device, and was the only noticeable app. This lowered confusion a lot with finding the app, and realising where to click. If a coach by accident clicked the home button, they immediately found their way back.

In the future, for new coaches, training how to use a smartphone is needed, before they are handed the device. While the YoungDrive is now simple to use to maybe not need an introduction, it would still help how to act in the app (for example encouraging to be honest with answering "Are you sure?").

7.3 How can Test Questions be Developed to Support Entrepreneurship Learning?

A lot has been learned about formulating and designing multiple-choice quizzes in a digital format. Below, suggestions for improvements are presented.

7.3.1 Designing for honesty with "Are you sure?"

One idea during the ideation of iteration 3, was to give the student two scores: one showing how correct they were, and one with how confident they were. You can be correct, but still not be empowered (you are unsure but correct). Similarly, you can be incorrect but confident (not empowered).

The reason for not having a "empowerment bar" (combining correct and sure, giving a summative score), was given by Josefina from YoungDrive: "The coaches

might want to game the system to get a better score, or be confused by how they got their score". For this reason, the coaches have stated they accept getting minus points if the are sure but incorrect, which is easily understood by them and feels fair.

In the app now, there is still a struggle with coaches answering positive on "Are you sure?", even if they are not. Reasons are among others that they say they are *partly* sure, and sometimes that they think they will be more punished for not being sure *than* being sure and incorrect. As this is not the view of the teacher, this needs to become much more clear in the app. One idea is to ask "How sure are you?", but keep the binary scale of only two answers. This would mean, that the coach needs to think about *how* sure she is about the answer, instead of *if* she is sure, which could have metacognitive benefits.

7.3.2 Self-Reflection After a Youth Session

When discussing the goals for iteration 3, Josefina talks about a need she has noticed during the coaches' roll-out in Zambia, where the app could help: doing self-reflection after a youth session. She says that this is at least as important as the coach training, especially in cases where Josefina or other project leaders don't have the resources to visit the coaches physically.

It is determined that while physical follow-up meetings are essential, the app can be used to help the coach doing a structured self-assessment and self-reflection exercise. When meeting with the teacher, this data could guide the coach-teacher discussion, the teacher focusing on discussing the most interesting data. This does not need to be a new app. Questions can be asked in a way that they are indeed meta-cognitive, encouraging learning by reflection.

Josefina mentions that when she is there to give feedback, it is very clear to the coach that he or she lacks knowledge and has not prepared enough. An app with self-evaluation and monitoring, would help keep the coach thoughtful and give the coach important insights. They are described to sometimes over-estimate their own knowledge.

7.3.3 Avoiding Memorization of Answers

To avoid memorization, the alternatives should be randomized in the future. While it is unlikely that a coach has an easier time remembering the correct answer by order instead of content, since they only repeats the wrong answers until the certification test, it is an extra measure.

7.3.4 Improving the Questions

Data analysis of results on specific questions could give a lot of insight, both into coach behaviour, and misunderstandings of questions, in the future. Here, a lot of data is already collected to be able to guide conclusions. Not only are questions recorded with correct and if answered confidently, but also number of tries per coach, and if it is a training quiz or certification quiz.

Simple analysis could be for example mean seeing what are the most difficult questions, where most people have answered wrong repeatedly. From the interviews in iteration 4, it is explained that some answers might be answered wrong because of for example difficult wording of questions, not necessarily because of lack of knowledge. To avoid this, data analysis could be effective, together with getting input from the teacher and coaches.

Improving the questions today has mostly been made from direct feedback from the coaches, or comparing quiz question formulation with current and desired level on Bloom's revised taxonomy. Regarding mapping educational objectives, it needs to be made sure that there are questions for each educational objective of the topic, which has not been done today. In Yengin and Uzanboylu (2012) questions were designed to support gradual knowledge building with an alignment to Bloom's Taxonomy, which could also be a viable alternative for this app, where questions are currently formulated as information appears in the participant and coach manual respectively.

7.4 How does Design Affect Usability and Learning Done via the App?

There is a clear connection between design and usability and learning. Here, future work for how design would improve usability or learning is presented.

7.4.1 Assessing Coach Guide Knowledge Before the Youth Session

When asked about the Zambia coach rollout, Josefina points out several challenges. "It feels like some of the coaches forgets the coach guide, even if it has been improved and better integrated with the participant manual. Some of them, don't even use the coach guide." This speaks for that the app should include quizzes for all coach guides as well. When asked if the coach guide quiz are more important than the topic quizzes, she answers that the correct knowledge is more important, because that is the one that needs to be explained correctly to the youth. Therefore, it should be moved into Future work.

7.4.2 Using a Flashcards Approach

In the ideation for iteration 2, flashcards are discussed again, with Henrik Marklund. In iteration 3, this was tested as a lo-fi material with successful results, but more work should be done.

In the ideation of iteration 4, a proposal was given that did not have time to implement. Therefore, the idea is described here: At the coach's second quiz try (having assessed and reflected on the knowledge), flashcards could be introduced to assist the coach in retrieving from memory, before getting the multiple-choice. For future work, when in Training after the first quiz try, The question should be shown *before* the answers are shown, and prompt the coach to think aloud about

what they think the answer is, before receiving the alternatives. The coach might be hindered from progressing to the multiple-choice answers until the app has understood the coach has thought hard about their answer to the question.

This is a good use of scaffolding, slowly introducing complicated app features. The hypothesis is inspired from Bjork (2016), that knowledge is strengthened if the coach retrieves from memory, versus looking up the answer or choosing the most likely answer.

7.4.3 Adding more media channels to more closely simulate the learning environment

To simulate the entrepreneur coach environment more accurately could possibly increase memorization of procedural learning objectives. Research exists that supports how using multiple channels like audio, video, voice could be beneficial, or to use interactive simulations. The advantage of multiple-choice from a developer perspective is that data can be collected easily and because of ease of implementation. From YoungDrive's perspective, it serves the target group of the coaches being first-time smartphone users well (Björling and Lönn, 2016).

7.4.4 Improvements to the Certification Mode

Following the advice from Sierra (2015) of quickly giving the user a feeling of a superpower, this should be becoming a Certified coach in the future. From the end results of iteration 4, we can learn that notably the intrinsic motivation is high, deliberate practice is present, and the coach can feel the intrinsic reward of having pushed herself and learned the material when certified. This is very positive.

This reaction, could and should be even more amplified when certified. It is discussable if this should be done by simple gamification, but an opinion by a coach was that medals earned should be more visible and that sounds could strengthen the feeling of achievement. Also, the quiz list could show these results, increasing motivation to take other quizzes that you have not yet mastered, or to better your score in a topic where you had not become certified.

7.4.5 Improvements training Correct Structure and Time Management

During all app tests (iteration 2-4), it has been shown that since Correct Structure and Time Management are both ordinal, the Training mode for such topics would be more suitable as interactive exercises than multiple-choice. The proposal is to first use drag-and-drop to place each activity of a youth session in the correct order, and then selecting the right time for the each activity. This assists the coach in creating a mental model, which can be used to retrieve from memory during the assessment.

7.4.6 Scaffolding with Flashcards

After the coach's first new try, Flashcards could be introduced to assist the coach in retrieving in memory, before getting the multiple-choice. To do this after the first assessment, is partly because of technology scaffolding (introduce new concepts in steps), partly because the knowledge is strengthened if the coach retrieves from memory versus looking up the answer or choosing the most likely answer (Bjork, 2016).

7.4.7 Memory Design

For the ideation of iteration 4, it was pointed out by pedagogical expert Henrik Marklund that if knowledge is to be memorized, memory techniques could be used. One such e-learning tool is Muth (2016). The tool has interactive learning modes, aiming to learn facts and terms with speed. This was not prioritized because of time constraints working with technical features that were not essential. Also, the idea was never proposed by users, only by experts. Moreover, the teacher opposed the idea of remembering answers that were not in the factual remember category. To do so, would oppose the learning objectives, which score higher on Bloom's revised taxonomy. However, to study how the coaches can remember better via an app, and learn memory techniques via the app, could be a future work which is advisable.

7.4.8 Sharing with One Another

In future versions of the app, mechanics of sharing content and co-creation would add value connected to Bloom's, reaching Create and Apply. Adding these game elements goes in line with Clark's research, which showed a positive correlation with learning and games that required multi-player collaboration (Clark and Killingsworth, 2014).

7.4.9 Including the Paper Manuals in the App

Already in iteration 2, it was proposed by coaches if the participant manual and coach guide could be included in the app, instead of as paper. The teacher agreed in principle, especially as the paper coach guide is often overlooked (reading the digital version could be designed to be made mandatory before unlocking other features), but also saw several challenges. Because of broadband limitations the manuals would take a lot of space to download. Also, the manuals are designed in A4 format, while a smartphone screen or even tablet is much smaller, making it cumbersome to read the whole manuals in the same way you would with the paper versions.

However, for the ideation of iteration 3, it was realized that if the manuals could be converted into smaller chunks, there is an opportunity for bite-sized spaced learning, instead of massed learning. How to split the manuals into these small parts? Well, by extracting the important parts for understanding an answer to a specific question, see figure 7.2. The teacher emphasised that this would take

time for her to do, but that the effort would be worth it. While there were no time for including the whole manuals, there was an idea the teacher thought acceptable for iteration 4: including on which page the coach would find the answer to a specific question. This was not one of the most prioritized features of iteration 4, which it was never done, but for further work, this should be investigated.

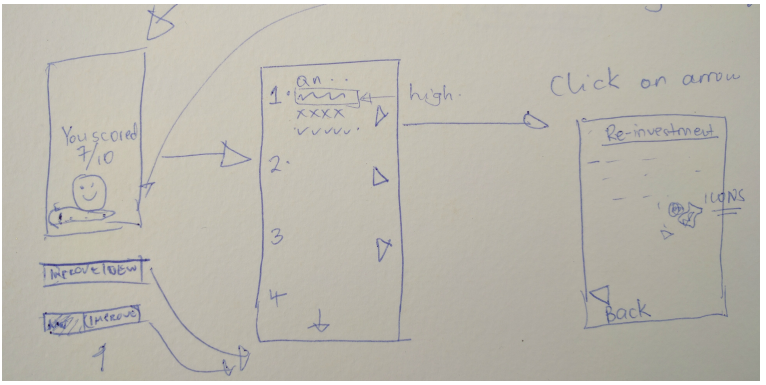

Figure 7.2: *By clicking on an arrow next to the question showing the correct answer compared with your answer, the coach could read an extract from the YoungDrive manuals, which includes the right answer. The learning benefit compared to just observing the right answer, is that the coach gets the answer in context, which improves memorization and understanding. An idea would be to no longer give the coach the correct answer in the score board, but to instead let the coach choose the right answer after reading the text.*

A major opportunity doing this would be to replace the paper manuals. Today, the manuals are the most expensive post of YoungDrive, which is why this would be attractive. While a smartphone or tablet could be returned after a YoungDrive program has ended, this is not possible with the paper versions, so there might be a monetary incentive to do so. A problem that would need to be addressed is that the manuals also have written exercises in them. These exercises could be made interactive in the app, which would also allowed for smart features such as automatic assessment, financial literacy simulations, etcetera. While this was not a scope of the master thesis, it is interesting further work to test if such interactive exercises and simulations would be effective in the developing country context of learning and teaching entrepreneurship.

7.5 How can Users' Feedback be Used to Inform Modifications of the App?

The users themselves has been essential in creating a valuable app. Users include coaches, but also the teachers and stakeholders. Below, one of these aspects are presented together with future work to answer the research question.

7.5.1 Educator Dashboard

Josefina has no means of accessing live quiz results today, as there is no educator dashboard developed. Instead, quiz results today needs to be transferred from a database into Google Sheets, which is cumbersome and not user-friendly.

There was not enough time to develop an educator dashboard, even if this had been a goal. Instead, low-fi trigger material was created, and co-creation stakeholder workshops were held, both for iteration 3 and 4.

In the future, this will be a must-have, and it ties well into YoungDrive's future wish of strengthening its quality assurance via monitoring and evaluation. Exactly how powerful the educator dashboard should be, might turn into a ethical discussion. One could argue that measures needs to be taken so that the app focus is to keep supporting the coaches to become better, and not to punish those coaches that doesn't have the same quiz results as others. As the combination of "Are you sure?" and correctness can give insight into the attitudes and care of the coaches, carefulness must be taken.

From an empowerment perspective, the Educator Dashboard can guide both the teacher in her teaching, allow her to have better coach feedback sessions, and doing future improvements of the YoungDrive program.

7.5.2 To Answer the Research Question More Elegantly

Working with user feedback to inform modifications of the app should be as central as has been during the master thesis. To further answer the research question, new situations and methodology could be tested. The new methodology Digital Service Design has been a success, and could be shared to other organizations than YoungDrive, especially (but not exclusively to) when working with technology in other developing countries.

Future work should also include to look at research suggesting how to work with balancing user feedback, stakeholder opinions and research. Such findings would be especially relevant when prioritizing and working the the backlog and doing sprint planning.

When it comes to user feedback, it comes in various ways: both qualitative data like interviews, and quantitative data such as quiz results, could be seen as a form of feedback. This is evident by the large number of methods used in the master thesis, see section 3.5 Data Analysis Framework. As such, it would be interesting to measure how the app is used after this finished master thesis, and if the usage of the app suggests what additional work needs to be done. The main output of the app should be better-performing coaches, and as long as the coaches are part of and feels ownership of the product development, the app will continue to give great results for the youth sessions.

Bibliography

AAU, DRT, UNNGOF. *Lost Opportunity? Gaps in Youth Policy and Programming in Uganda*. ActionAid International Uganda, 2012.

S. K. Abell and N. G. Lederman. *Handbook of Research on Science Education*. Psychology Press, 2007.

Agile Alliance. The agile manifesto. `http://www.agilemanifesto.org/`, 2001.

Apache. Apache cordova. `https://cordova.apache.org/`, 2011.

Habanero Quiz Apps. Quizoid. `https://play.google.com/store/apps/details?id=de.habanero.quizoid&hl=en`, 2016.

Quian S. Miao C. Bae, T. J. and J. O. Fiet. The relationship between entrepreneurship education and entrepreneurial intentions: A meta-analytical review. *Entrepreneurship Theory and Practice*, 38(2), 2014.

Mushomi J. A. Nahalamba S. Byakagaba J. Magero S. Sheila Ariko J. O. Balyesiima, D. F. and R. Kaggwa. *The State of Uganda Population Report 2012*. The Republic of Uganda, 2012.

R. Bjork. Applying cognitive psychology to enhance educational practice. `http://bjorklab.psych.ucla.edu/research.html`, 2016.

I. Björling and J. Lönn. *Exjobbs-resultat: Coach-tränings-app*. YoungDrive, 2016.

Heer J. Ogievetsky V. Bostock, M. D3.js - data-driven documents. `http://ischool.zm/`, 2014.

R. O. Brinkerhoff. Making l&d matter. `https://www.youtube.com/watch?v=__wUZzJV8lM`, 2013.

F. Brühlmann. Gamification from the perspective of self-determination theory and flow. Master's thesis, University of Basel, Switzerland, 2013.

K. Chang. Parallel coordinates toolkit : Parcoords.js. `https://syntagmatic.github.io/parallel-coordinates/`, 2012.

Bruno V. Cheong, C. and F. Cheong. Designing a mobile-app-based collaborative learning system. *Journal of Information Technology Education: Innovations in Practice*, 11, 2012.

Tanner-Smith E. Clark, D. and S. Killingsworth. *Digital Games, Design and Learning: A Systematic Review and Meta-Analysis (Executive Summary)*. SRI International, 2014.

S. Clatworthy. Interaction design: services as a series of interactions. In M. Stickdorn and J. Schneider, editors, *This is Service Design Thinking: Basics, Tools, Cases*, chapter 11, pages 80–87. Consortium Book Sales and Dist, 2010.

Drifty Co. Ionic framework. `http://ionicframework.com/`, 2016.

Hilera J. R. Burchino R. Jiménez L. Martínez J. J. Gutiérrez J. A. Gutiérrez J. M. de Marcos, L. and S. Otón. An experiment for improving students performance in secondary and tertiary education by means of m-learning autoassessment. *Computers and Education: An International Journal*, 55(1069-1079), 2010.

E. L. Deci and R. M. Ryan. *Intrinsic motivation and selfdetermination in human behavior*. Plenum, 1985.

E. L. Deci and R. M. Ryan. Intrinsic and extrinsic motivations: Classical definitions and new directions. *Contemporary Educational Psychology*, 25(54-67), 2000.

Solomon G. T. Dickson, P. H. and K. M. Weaver. Entrepreneurial selection and success: Does education matter? *Journal on Small Business and Enterprise Development*, 15(239-258), 2008.

J. Dirksen. *Design for How People Learn*. New Riders, 2012.

Duolingo. Duolingo. `https://play.google.com/store/apps/details?id=com.duolingo&hl=en`, 2016.

C. Dweck. The power of believing that you can improve. `https://www.ted.com/talks/carol_dweck_the_power_of_believing_that_you_can_improve?language=en`, 2014.

J. S. Eccles and A. Wigfield. Motivational beliefs, values, and goals. *Annual Review of Psychology*, 53(109-132), 2002.

A. J. Elliot and M. V. Covington. Approach and avoidance motivation. *Educational Psychology Review*, 13(2), 2001.

Instagram Facebook. React. `https://facebook.github.io/react/`, 2016.

Grameen Foundation. Ledger link. `http://awards.ixda.org/entry/2014/ledger-link/`, 2014.

S. M. Fulmer and J. C. Frijters. A review of self-report and alternative approaches in the measurement of student motivation. *Educational Psychology Review*, 21(219-246), 2009.

Inc. GitHub. Github. https://github.com/, 2008.

R. Heer. A model of learning objectives. http://www.celt.iastate.edu/wp-content/uploads/2015/09/RevisedBloomsHandout-1.pdf, 2012.

Kolvereid L. Iakovleva, T. and U. Stehphan. Entrepreneurial intentions in developing and developed countries. *Education + Training*, 53(5), 2011.

Meteor Development Group Inc. Meteor. https://www.meteor.com/, 2016.

iSchool. ischool zambia. http://ischool.zm/, 2014.

H. Kniberg. *Scrum and XP from the Trenches - 2nd Edition*. InfoHQ, 2005.

T. R. Koballa and S. M. Glynn. *Attitudinal and motivational constructs in science learning*. Focal Press, 2007.

D. R. Krathwohl. A revision of bloom's taxonomy: An overview. *Theory Into Practice*, 41(4), 2002.

D. F. Kuratko. The emergence of entrepreneurship education: Development, trends, and challenges. *Entrepreneurship theory and practice*, 25(5), 2005.

Meister Labs. Mindmeister. https://www.mindmeister.com/, 2016.

H. Lopez. Parallel coordinates: Read out patterns. http://une-terre.blogspot.se/2012/09/parallel-coordinates-read-out-patterns.html, 2016.

Bligh B. Manches A. Ainsworth S. Crook C. Luckin, R. and R. Noss. *Decoding Learning: The Proof, Promise and Potential of Digital Education*. Nesta Operating Company, 2012.

J. Lönn. Youngdrive zambia coach interviews. , 2015a.

J. Lönn. Youngdrive zambia coach interviews, 2015b.

J. Lönn. Youngdrive uganda: Statistik youth, cbts, projektledare - 2015, 2015c.

J. Lönn. *YoungDrive Participant Business Booklet*. YoungDrive, 2016.

J. Löwgren and E. Stolterman. *Thoughtful Interaction Design: A Design Perspective on Information Technology*. The MIT Press, 2007.

Busenitz L. W. Bird B. Marie Gaglio C. McMullen J. S. Morse E. A. Mitchell, R. K. and J. B. Smith. The central question in entrepreneurial cognition research. *Entrepreneurship Theory and Practice*, 31(1), 2007.

VOTO Mobile. Voto mobile. https://www.votomobile.org/, 2012.

C. Muth. Memorize. http://memorize.com/, 2016.

D. Nicol. E-assessment by design: using multiple-choice tests to good effect. *Journal of Further and Higher Education*, 31(1), 2007.

S. Nissar. Social innovation and entrepreneurship in uganda. why mobile services are growing fast in the area. http://bit.ly/medium-article-expedition-mondial, 2016.

M. J. I. Oviawe. Repositioning nigerian youths for economic empowerment through entrepreneurship education. *European Journal of Educational Studies*, 2(2), 2010.

D.H. Pink. *Drive: The Surprising Truth about what Motivates Us*. Canongate, 2011. ISBN 9781847677693.

L. Pittaway and J. Cope. Entrepreneurship education a systematic review of the evidence. *International Small Business Journal*, 25(5), 2007.

T. Ropinski. Tnm067 - scientific visualization: Chapter 1: Basics, 2014.

E. Ruskovaara and T. Pihkala. Entrepreneurship education in schools: Empirical evidence on the teacher's role. *The Journal of Educational Research*, 108:3: 236–249, 2015. doi: 10.1080/00220671.2013.878301.

Salesforce. Heroku. https://www.heroku.com/, 2007.

K. Sierra. *Badass: Making Users Awesome*. O'Reilly Media, 2015. ISBN 9781491919071.

Bell B. S. Kraiger K. Sitzmann, T. and A. M. Kanar. A multilevel analysis of the effect of prompting self-regulation in technology-delivered instruction. *CAHRS Working Paper Series*, 8, 2008.

Gino F. Pisano G. P. Stefano, G. D. and B. R. Staats. Learning by thinking: Overcoming the bias for action through reflection. *Harvard Business School Technology and Operations Management Unit Research Paper Series*, 14(093), 2015.

M. Stickdorn and J. Schneider. *This is Service Design Thinking: Basics, Tools, Cases*. Consortium Book Sales and Dist, 2010. ISBN 9789063692568.

Ayres P. Sweller, J. and S. Kalyuga. *Cognitive Load Theory*. Springer-Verlag New York, 2011. ISBN 9781441981264.

Agaze Dev Team. Magic json. https://chrome.google.com/webstore/detail/magic-json/cajifcebjiflndefndbnoeenjpiiiagm?hl=en, 2016.

W. B. Walstad and M. L. Kourilsky. *Entrepreneurial Attitudes and Knowledge of Black Youth*. Kauffman Center for Entrepreneurial Leadership, 1998.

E. Widmark and S. Nissar. Service design - en crash course. Expedition Mondial, 2016.

A. Wigfield and J. S. Eccles. Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, 25(1), 2000.

Ince I. F. Karahoca A. Karahoca D. Yengin, I. and H. Uzanboylu. The use of delib-erates practices on a mobile learning environment. *1*, (241-249), 2012.

YoungDrive. Youngdrive - website. `http://youngdrive.org/`, 2016.

Appendix

A

Appendix A: Images of Developed Application

In the following appendix, the graphics from the final app (iteration 4) is shown in detail according to the figure from section 4.1 Developed Application. The flow is choosing a quiz (figure A.1), logging in (figure A.2), taking the quiz (figure A.3), getting feedback (figure A.4), redoing the quiz (figure A.5), being able to get certified when 100% correct (figure A.6), testing if you can now be 100% correct in one try, and if passing, getting the certification (figure A.7).

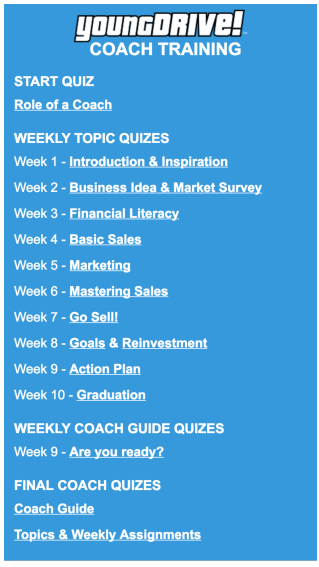

Figure A.1: *Choose quiz*

Figure A.2: *Login*

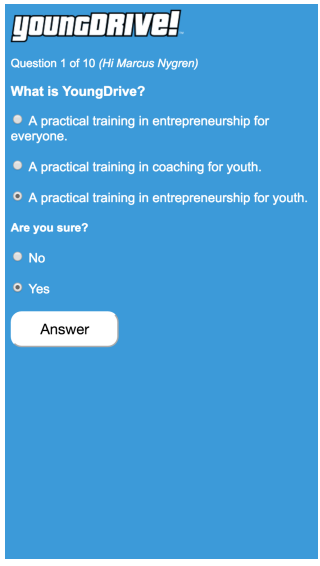

Figure A.3: *Self-assess*

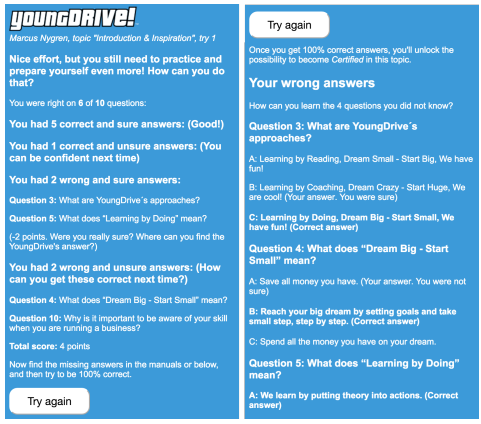

Figure A.4: *Get feedback and try again*

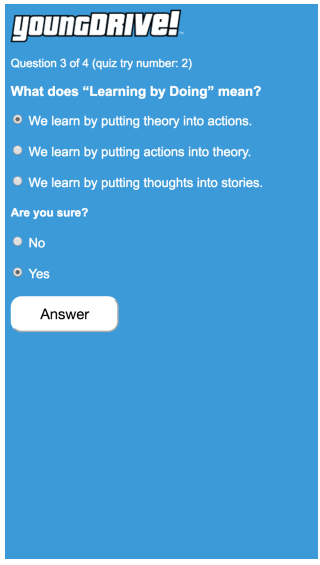

Figure A.5: *Train until 100% correct*

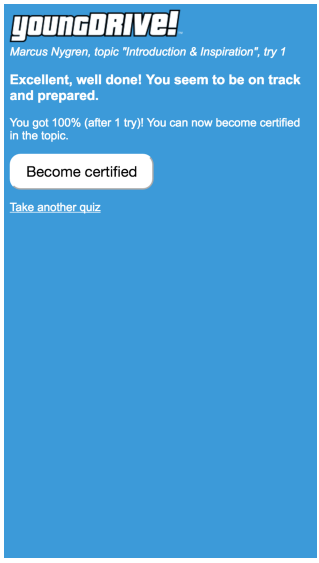

Figure A.6: *Unlock the possibility to get certified when 100%*

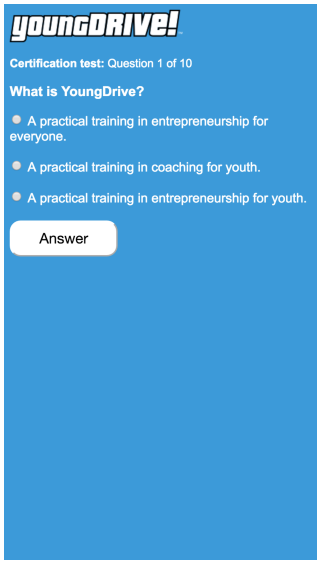

Figure A.7: *Test if you can get 100% in one try*

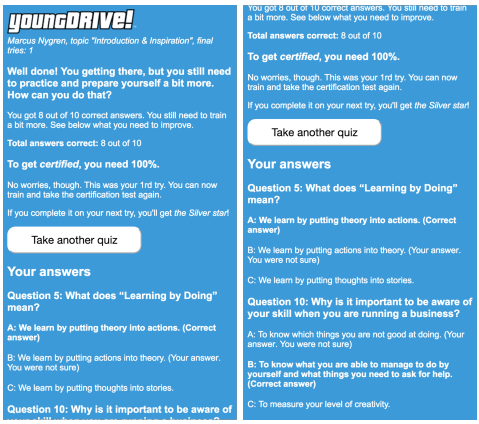

Figure A.8: *If not, get feedback and go back to training*

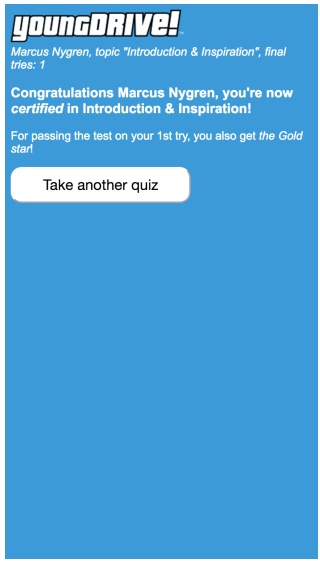

Figure A.9: *If 100% correct, you get certified in the topic*

B

Appendix B: Pre-Test

In the following appendix, the pre-test used for iteration 4 is shown, see figure B.1 and figure B.2. The pretest is based on Walstad's questionnaire, with modifications to be understood in the Tororo context Walstad and Kourilsky (1998). The correct answers are 1) 2, 2) 1, 3) 2, 4) 3, 5) 2, 6) 2 and 7) 1.

Your name:

YoungDrive Pre-knowledge

School level? _____ **Age?** ____ **Gender?** Man Woman
Running own company? Yes No
Used to smartphone? Beginner Medium Experienced

Hours of preparation for 1 youth session? ___ hours
When? A few days before The day before The same day
How many occasions do you prepare for 1 youth session? 0 1 2 3

1. Which of the following best describes an entrepreneur?
   1. A manager of a large corporation
   2. A person who owns a small business
   3. A government official running a regulatory agency
   4. I think some, a combination, or all of the answers are correct
   5. I don't know

2. Over the last ten years, which of the following groups has created the most new jobs in the economy?
   1. Small business
   2. Large businesses
   3. The Federal government
   4. I think some, a combination, or all of the answers are correct
   5. I don't know

3. Considering the methods of raising capital to start a new business, which of the following is the most typical?
   1. Borrowing money from the government
   2. Using personal money or borrowing from friends or relatives
   3. Borrowing money from a bank
   4. Issuing company stock or bonds to the general public
   5. I think some, a combination, or all of the answers are correct
   6. I don't know

1

Figure B.1: *Pre-test, page 1, based on Walstad Walstad and Kourilsky (1998).*

Your name:

4. Which factor is most important for business survival?
   1. The value of the company's common stock
   2. Having a low depreciation rate
   3. That the company is profitable
   4. Having a board of directors
   5. I think some, a combination, or all of the answers are correct
   6. I don't know

5. Which one of three businesses is the best example of a franchise?
   1. The CocaCola corporation
   2. The Shell gas station
   3. The local cement factory
   4. I think some, a combination, or all of the answers are correct
   5. I don't know

6. The prices of most products in a competitive market like Uganda, are determined by the:
   1. The Consumer Price Index
   2. Supply and demand for products
   3. Local, state, or Federal government
   4. Monetary policy of the Federal government
   5. I think some, a combination, or all of the answers are correct
   6. I don't know

7. Which of the following do you think is the basic purpose of profits in a free market economy?
   1. Reward businesses for producing what consumers want
   2. Pay for the wages and salaries of workers
   3. Transfer income to the wealthy
   4. None
   5. I don't know

2

Figure B.2: *Pre-test, page 2, based on Walstad Walstad and Kourilsky (1998).*

C

Appendix C: Time Plans and Activities

In appendix 1, the original time plan is presented together with the revised time plan, done for the half-time evaluation with the examiner.

C.1 Original Time Plan and Activities

C.1.1 Before Uganda

Week	Focus
2	Workshop with Lena Tibell and Konrad Schönborn on Research questions & Proposal of method.
3	Start writing "Planeringsrapport". Study interaction design via guest lecture Jonas Löwgren, and reading the book "Thoughtful Interaction Design".
4	Interview with Take Aanstoot, Social entrepreneur in Kenya. Submission "Planeringsrapport". Education day in Service design in Stockholm (by Expedition Mondial). Meet Joachim Svärdh about Entrepreneurship research.
5	Approval "Planeringsrapport" with Camilla Forsell. Meeting with Lena Tibell and Konrad Schönborn (2016-02-02). Travel to Uganda.

C.1.2 In Uganda

Times specified are in local time to where the master thesis was done at the time. Uganda time (EAT - Eastern Africa Time) is 2 hours forward of Swedish time (CET - Central European Time). Meetings with Swedish partners are generally done via Skype, where Uganda meetings are preferably done in person. Note

that during all of this time, writing the master thesis will progress. After the time in Uganda, the report will be a 100% focus. 1 day per week will be spent on report writing, including Analysis work for the meetings.

Week	Focus
6	**Cultural adaption**. Land, set up wifi, set up the apartment, learn about the YoungDrive organization, meet people. Be prepared for stomach disease. Get familiar with the transportation system in Kampala. Get familiar with the city. **Iteration 1**. Prepare Iteration 1 with Iliana. Start-up meeting with partners. Start report writing: analyze, collect material, sort, structure and plan.
7	**Iteration 1**. Prepare Interactions. Analyze Start-up meeting with partners. Write on report. in order to create *Questionairee guide*. Understand technical tools, without working on an app solution - the goal is to get familiar with the tools.
8	**Iteration 1**. Travel for Interactions. Do 8 face-to-face interviews, with no digital focus, hypotethical situations. Do minimum 2 field visits to understand the coach's situation, ideally living in Kamuli or Tororo a couple of days. This is a good opportunity to learn coaches how the tables and smartphones work.
9	**Iteration 1**. Analysis & Compilation. Thursday: Expert meeting (March 3rd, 6-7 PM). Friday: Partner meeting (March 4th, 11-12 AM). **Iteration 2**. Determine Needs. Ideation. Create low-fi Trigger material (pen and paper) and determine what the hi-fi (digital app) material should be.
10	**Iteration 2**. Design and Develop the hi-fi trigger material. *Half-time check-up with examiner.*
11	**Iteration 2**. Interactions, control group #1 & #2.
12	**Iteration 2**. Interactions, control group #1 & #2.
13	**Vacation**.
14	**Iteration 2**. *Analysis #2* (What choices needs to be made? What path should be taken? Start formulate Customer path. If needed, document how people see apps, document limitations, document experience needs, document risks.) & Compilation. Thursday: Expert meeting (April 7th, 4 PM). Friday: Partner meeting (April 8th, 11-12 AM). Continued Development Creative Brief. Determine what actions needs to be taken outside of the development of the app. Create Behovsgrupper.
15	**Iteration 3**. Develop and Modifications phase.
16	**Iteration 3**. Develop and Modifications phase. Interactions: App Tests with Interviews & Measurements (with time allocated for late arrivals and missing participants).
17	**Iteration 3**. Interactions: App Tests with Interviews & Measurements. Analysis & Compilation. Friday: Partner meeting (April 29th, 11 AM) & Expert meeting (April 29th, 4 PM).
18	**Final analysis**. Finalize the app. Travel back to Sweden.

C.1.3 After Uganda

Week	Focus
19	Write on Master thesis report. Attend Auscultations.
20	Write on Master thesis report. Attend Auscultations.
21	Write on Master thesis report. Attend Auscultations. Find opponent for Master thesis.
22	Submission of report to examiner, after approval by supervisor. Examiner decides on date and time for presentation. Send report to opponent, and get the opponent's report.

C.1.4 After Semester

Week	Focus
35	Presentation of Master thesis, with supervisor, examiner and opponent. Hand over publication approval to the administrator.
36	Opposition of another person's Master thesis.
37	Do changes to report if requested. Upload report to X-sys for approval (within 10 days). Write Reflections document and submit on X-sys within the 10 days. Publish master thesis in X-sys.

C.2 Revised Time Plan and Activities

When the possibility of attending the YoungDrive Zambia training appeared, week 10-18 of the original time plan needed to be revised. These were the changes made, which were also approved by the examiner during the half-time evaluation of the master thesis:

Week	Focus
10	Iteration 2 Trigger material & Prepare Interactions
11	Iteration 2: Interactions (at YoungDrive Zambia)
12	Vacation in Uganda
13	Iteration 3: Insights
14	Iteration 3: Ideation & Trigger material
15	Iteration 3: Interactions (in Tororo) & Insights
16	Iteration 4: Ideation & Trigger material
17	Iteration 4: Interactions (in Tororo) & Insights
18	Final analysis. Finalize the app.

Below, the activities for the interactions of each iteration is presented.

C.2.1 Iteration 1: Uganda Coach Visit

In iteration 1, the following activities were carried out and then analysed:

- Stakeholder interview (for Need groups)

- Coach interview and field visit (for Need groups)
- Workshop (for Customer Journey Map)
- Smartphone test (for Need groups)

C.2.2 Iteration 2: Zambia Coach Training

In iteration 2, the following activities were carried out and then analysed:

- App test observations (for Sprint backlog)
- Quiz results (for Sprint demo)
- App design workshop: use case 1 (for Sprint backlog)
- App design workshop: use case 2 (for Need groups)

C.2.3 Iteration 3: Uganda Formative Test

In iteration 3, the following activities were carried out and then analysed:

- App test observations (for Sprint backlog)
- Service Mini-Sprint - 5 workshops (for Sprint backlog)
- Field visits (for Sprint demo)
- Small formative app test (for Sprint backlog)
- Big formative app test (for Sprint demo)

C.2.4 Iteration 4: Uganda Summative Test

In iteration 4, the following activities were carried out and then analysed:

- Big summative app test (for Sprint demo and quantitative analysis)