Exploring Motivation and Engagement in Digital Artefacts : a Self-Determination Theory Perspective

Marieh Sayadchi

Master's thesis

Supervisor:

David Lamas, PhD

Tallinn University

School of Digital Technologies

Tallinn 2016

Authors Declaration

I, Marieh Sayadchi, author of this thesis, declare that the thesis concerned is the result of my own independent research and it has not been previously submitted for a defence.

Non-exclusive Licence

To reproduce a thesis and make thesis available to public

I, Marieh Sayadchi (date of birth: 19/06/1985)

1. grant Tallinn University a permit (a non-exclusive licence) to reproduce for free and make public in the repository of Tallinn University Academic Library a piece of work created by me

"Exploring Motivation and Engagement in Digital Artefacts : a Self-Determination Theory Perspective"

supervised by David Lamas.

2. I am aware of the fact that the author also retains the rights mentioned in Clause 1.

3. I certify that granting the non-exclusive licence does not infringe the intellectual property rights of other persons or the rights arising from the Personal Data Protection Act.

In Tallinn, (digital) signature and date

Abstract

With the wide-spread use of digital applications in different aspects of people's lives, the importance of supporting motivation in digital artefact design is getting more prominent. Self-Determination Theory (SDT) as a widely used theory of motivation has shown promising explanatory power in several other areas such as video games. The relevance of exploring theories from psychology in Human-Computer Interaction (HCI) is that it can inform designers throughout the design process. When proved applicable in this new context, related methods and scales can also be used in different steps of design and evaluation.

The purpose of current study is to explore Self-Determination Theory as a possible theoretical background for designing for motivation and engagement in digital artefacts. This thesis investigates the possible use of SDT in general digital design as it has been used in other fields. To provide empirical data one laboratory experiment and one online study have been carried out and are described in this work. Both studies show promising results indicating that SDT can explain motivation and engagement in digital artefact design hence, it is worth exploring further.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Perhaps we all can agree on the importance of motivation in performing tasks and pursuing goals. From mechanistic period in 1930s to more contemporary attempts, psychologists have studied motivation, and conducted experiments to understand this phenomena for decades(for more history of motivation studies, see (Graham & Weiner, 1996)). Among the branches of motivational theories, the ones which discuss external and environmental factors are especially insightful for practical implementations. These external factors can potentially help or hinder the formation and level of the motivation. Therefore such theories have the inherent possibility to guide the design. Their guidance can lead the way for changing and improving conditions in order to facilitate motivation and engagement.

Since its formation, Self-Determination Theory (SDT) and its branches have been used in that sense, to inform educators, parents, caregivers and employers in designing and planning conditions to maximize motivation in their target groups. Many classroom experiments have been conducted based

on this theory to inform better lesson planning and teacher-student interactions in pursuit of achieving higher motivation and deeper learning. Early studies in classrooms included for example effects of positive and negative performance feedbacks in enhancing or diminishing intrinsic motivation (e.g., Deci, 1971 also Deci & Cascio, 1972), autonomous versus controlling class management by teachers and their outcomes on students' optimal learning and intrinsic motivation(E. A. Skinner & Belmont, 1993). In a more recent research Yaniv Kanat-Maymon et al. studied academic dishonesty and the role of SDT's basic needs deprivation in increasing the likelihood of academic deception(Kanat-Maymon, Benjamin, Stavsky, Shoshani, & Roth, 2015).

In companies and workplaces SDT is used to motivate employees in their jobs, resulting in higher performance and better mental state and job satisfaction. (Gagné & Deci, 2005) discusses the controversial role of tangible reward in work place and its effect on employees, as well as how to autonomize extrinsic motivation. They also suggest that "intrinsic motivation (based in interest) and autonomous extrinsic motivation (based in importance) are both related to performance, satisfaction, trust, and well-being in the workplace"(Gagné & Deci, 2005).

Importance of designing digital tools, applications and digital environments in general, for motivation and engagement is getting more evident in recent years, since people spend more time than ever in digital world, working, studying and playing. When students regularly use virtual learning environments as part of their education, putting adequate thoughts on designing these tools to foster motivation becomes as important as planning lessons and managing classrooms. The same also applies for all the digital

applications to some extent.

The purpose of this study is to examine SDT as a possible theory to inform design for motivation and engagement. In absence of a theoretical backbone for motivational design, several trends have risen to help practitioners in design process. These trends often focus on the examples of the effective design rather than a theory that explains reasons why one design works and the other does not. Example based approach, although useful, can often become limiting, theory based design on the other hand opens a broader range of possibilities and has greater generative and explanatory powers. At the end it is optimal to have both in hand for a good design practice.

The first time SDT is used in empirical study in digital world was in 2006 by Ryan and colleagues, part of the core team in developing SDT itself (Ryan, Rigby, & Przybylski, 2006). Their empirical study focused on mapping SDT in video games and addressing the question of how this theory might explain the motivational pull of video games. They conducted 4 sub-study examining different hypotheses. Since then others have done similar empirical studies on games featuring different aspects of the phenomena. For example Tamborini et al. (2010) studied intuitive controls and relatedness in an experimental video game setting (Tamborini, Bowman, Eden, Grizzard, & Organ, 2010).

The explanatory power of SDT proves relevant in game context based on the previous studies, however little is done to bring self-determination theory(as a cluster of its branches), to other digital tools and applications.

In the theoretical and analytical level, this theory looks promising to help inform motivational design in areas other than game context as well.

In the fields other than games, gamification[1] being inspired from game design, has been linked to SDT. Gamification looks at games for inspiration in order to afford same interest and fun in a non-gaming context. yet, during my literature review I have not came across empirical studies of gamification to measure SDT related factors either.

Current work is inspired by the video game studies such as(Ryan et al., 2006) . Several others have implemented similar experiments on (digital) games as well. However, I will be repeating it in non-game applications, and make some changes that reflect this specific context and my personal research interests. In short, current study tries to find out if SDT can be used in general digital design in the same way that it has been used to study video games. Especially if the scales and methods prove relevant in this context, it can indicate that motivational designers can use the same approaches to : **a)**evaluate and compare designs **b)**utilize them in user tests during design and **c)**use SDT/its branches to inform their design and predict the outcomes.

1.1 Research Problem and Goal

The main aim of this work is to explore a theoretical background for designing for motivation and engagement. Motivational design is not a strong trend in Human-Computer Interaction(HCI) at the moment. It is mostly limited to specific use-cases such as education, behaviour change or health care. How-

[1]"Gamification refers to: a process of enhancing a service with affordances for gameful experiences in order to sup- port user's overall value creation" (Huotari & Hamari, 2012)or "the use of game design elements in non-game con- texts"(Deterding, 2012)

ever there is a great opportunity for general information technology design to benefit from psychological motivation theories such as self-determination theory. Motivation as defined within self-determination theory has various levels and types, therefore it can provide different guidances for each type of application. In short, SDT says that humans have three universal psychological needs: *competence*, *autonomy* and *relatedness*. Competence being the feeling of mastery, improvement and ability to perform well; Autonomy as feeling of choice and control over one's own actions and decisions; and relatedness the feeling of closeness to other people and belonging to a group or community(please refer to section 2.2 for more details and definitions on the theory). Designing while having fulfilment of these needs in mind gives rise to higher enjoyment and intrinsic motivation that in turn leads to continued use and satisfaction. In my master thesis I am exploring SDT, if it can explain motivation and engagement phenomena in general digital design and how it may inform design.

1.2 Research Question and Hypotheses

Based on the research goal explained in previous section, it is imperative to first find out if need satisfaction in this context will predict enjoyment and other positive outcomes as well(as it did in the context of video games). If the answer to this question is positive, it is the first step that sets the base for more specific studies in the future: The series of further studies with specific targets that will later help formation of a framework that can provide practical guidelines. Research questions 1, 2 and 3 address the applicability

of SDT in digital design and research question 4 tries to find some details on what features may be responsible for satisfaction(or thwart) of SDT based needs.

Research question 1: Does users' perceived need satisfaction(based on SDT) predict their using patterns of an application(App) such as frequency of use and how much they enjoy using the App?

Research question 2: Can SDT predict differences in popularity of two Apps?

Research question 3: What is the role of usability or intuitive controls (IC) in need satisfaction, does it differ in applications with different popularity levels?

In previous studies (Tamborini et al., 2010) and (Ryan et al., 2006) Intuitive Controls(IC) were included as a factor that might affect competence and autonomy perception of user. It is called "intuitive controls" in the context of video games, because a big part of the game being easy to use and intuitive is in its controls. These controls are used for moving characters around or performing actions in the game. The idea behind IC is to determine how intuitive a system is for user, how easy it is to learn and remember and sort out problems. Sample of the questions asked to determine IC was "Learning the game controls was easy." This measure is very close to the idea of usability in digital tools in general. Standard and commonly used scales of usability such as "system usability scale" are perhaps more suited to the conditions of current work and also cover more dimensions of usability.

Research question 4: Can most liked and disliked features of the apps be explained by SDT?

Based on the research questions above three hypotheses are developed. Hypothesis 1 relates to research question 1 and hypothesis 2 addresses research question 2 and hypothesis 3 is driven from research question 3. Research question 4 is addressed by interviews during the lab experiment and open-ended questions in online study.

Hypothesis 1 *User's perceived experience of need satisfaction in namely : competence, autonomy and relatedness, will predict three outcomes of : enjoyment, frequency of use willingness to use the application again and recommend it to others.*

This hypothesis main postulate is that each need being supported would have effects upon at least some of the outcomes examined. In other word this hypothesis says that in one application, need satisfaction will correlate with their enjoyment, frequency of their use or further use and recommendation to others, for its users.

Hypothesis 2 *There will be a significant difference between means of need satisfaction for three needs of : Competence, Autonomy and Relatedness, in two applications with difference in popularity. I expect the more popular application to have higher need satisfaction means.*

Hypothesis 3 *Does usability correlate with need satisfaction for Competence and Autonomy ? I hypothesize that higher System Usability Scale(SUS) value will predict higher need satisfaction.*

Side question 1: Does usability differ for applications ranked 1 and 10 ? In the study of (Ryan et al., 2006) they hypothesized and proved correlation between IC and experiences of autonomy and competence in game

play. As I discussed earlier intuitive controls as mentioned in previous studies, is equivalent of usability in digital design vocabulary. Thus there are more suitable questionnaires to measure usability such as System Usability Scale(SUS). These hypotheses and research questions seek to provide data driven evidence on applicability of SDT in designing applications. Still the experiments serves as small empirical evidence that will be followed by thorough analysis of each tool in light of the theory.

1.3 Some Definitions

In this section motivation and engagement will be defined based on the literature mostly related to HCI. Motivation from SDT point of view will be discussed under section 2.2.

1.3.1 Motivation

Motivation is widely referred as the force behind an action. What makes a person to do something and the will to change the current state and perform some action. Historically motivation has been defined in many ways(behaviour, goal etc.) concentrating the major ideas of the time.

1.3.2 Engagement

What is meant as engagement here is the attention and emotional involvement of user with the system and not broad sense of engaging in something as starting or doing something. This use of the word engagement somehow already implies a temporal duration as well, however short that duration

might be.

Some factors of engagement could be :

Attention: user is focused at the interaction/task in hand. User's attention on screen based systems can be measured as an example by seeing if she is looking at screen, her gaze on the relevant places of screen and of course the interactions between system and user.

Interest: in (Peters, Castellano, & de Freitas, 2009) interest is defined as importance and relevance of a task in order to engage the user. System's characteristics such as aesthetics, quality of interactions,level of challenges, controls, sensory appeal etc, are factors that can affect user engagement. From user's side factors like user's needs, goals and emotions will determine how user engages with the system(O'Brien, H.L. and Toms, E.G. , 2010).

Chapter 2

Background and Literature Review

In this chapter the relevant topics that were necessary to set the context and background of this work are described. Since the main idea behind current work is motivation and engagement in digital artefact design, first section overviews the different approaches that have been used in digital design for user motivation and engagement.

Sections 2.2 explains self-determination theory and its branches, which is vital for understanding the work.

The last section reviews the reported applications of SDT in digital world. It includes various sub-fields such as video games, educational tools and gamification.

2.1 Designing for Motivation and Engagement

Although designing for motivation and engagement is not so far a defined branch in human-computer interaction(HCI) by itself, concepts related to user engagement already appear in the field. Design for Motivation on the other hand, shows up mostly in educational digital design. Educational tools will be discussed under section 2.3, "SDT in digital world" . Engagement as attractive and engaging user interfaces have been discussed in several studies. Especially engagement through aesthetics and interface design, in the context of user experience, had the most attention in the past few years. *User engagement* in this sense has a restricted focuses on the quality of individual interactions, and not the long term experience throughout the whole lifespan of the product. While other user engagement theories such as flow theory(Csikszentmihalyi, 1990) and theories around arousal focus on a short term and one session use, SDT provides a wider and more general guides. Details on interactions and design elements can also be driven from SDT. But the main domain of its application and insight, is for longer term use of a system, and more fundamental in general. One example for that ,is the concept of "flow zone"(Sutcliffe, 2009) which is the situation that a task is neither too difficult nor too easy for the user, thus keeps the user involved. the same concept as flow zone can be driven from SDT based on competence need satisfaction as well. It is corresponding to "optimal challenge" in SDT competence discussions. The *flow* explains some aspects of engagement but does not cover all. O'Brien & Toms(2008) argue that although flow has a lot to do with engagement, there are differences. For

example, flow involves intrinsic motivation but engagement can happen in less intrinsically motivated activities(O'Brien & Toms , 2008). This is the place that evolution of SDT with its sub-theories come in to explain engagement in different levels and types of motivation(see organismic integration sub-theory 2.2.3). Although its difficult to argue that SDT fully covers flow theory, it explains the flow phenomena well enough and goes beyond as well.

2.2 SDT: A Short Review

Self-determination theory was first proposed in the period of psychological studies where operant theory(B. Skinner, 1971) was the dominant view in the field. In the first part of SDT Deci and Ryan suggested that people indeed have intrinsic motivation and engage in many things just because the task is fun and enjoyable for them as oppose to being always under influence of an external force. This theory says that intrinsic motivation exists by its own and apart from any external motivation. Such motivation can be observed in situations that people engage in activities because of curiosity, play, challenge etc. It is also different from actively and consciously seeking enjoyment or challenge, and acts in subconscious level.

In the following sections important vocabulary of this theory is defined and then sub-theories of SDT are introduced in short.

2.2.1 Some Definitions

Here are some definitions from the literature, that are most relevant to current work:

Motivation

Ryan et. al. defined a motivated person to be:

"To be motivated means to be moved to do something. A person who feels no impetus or inspiration to act is thus characterized as unmotivated, whereas someone who is energized or activated toward an end is considered motivated. "(Ryan & Deci, 2000)

by Lai's definition, motivation refers to:

"Motivation refers to reasons that underlie behavior that is characterized by willingness and volition. "(Lai, E. R. , 2011)

Intrinsic and Extrinsic Motivation

Defined by Ryan & Deci in their recent summary of SDT, intrinsic and extrinsic motivation differ in their inherent origins: "intrinsic motivation, which refers to doing something because it is inherently interesting or enjoyable, and extrinsic motivation, which refers to doing something because it leads to a separable outcome." (Ryan & Deci, 2000)

Although the name might be deceiving, even personal plans and goals are not considered "intrinsic motivation"(IM). In fact anything that is done for an outcome however personally planned and anticipated, is not intrinsic by this definition.

Competence

Competence is the feeling of being able to perform the task optimally,or being good at something.

Autonomy

The perception of choice and control over one's own actions called autonomy.

Relatedness

Feeling close to someone or something, feeling of belonging to a community or relating to them are senses of relatedness.

2.2.2 Cognitive Evaluation Theory

"Cognitive Evaluation Theory (CET) was presented by Deci and Ryan (1985) to specify the factors in social contexts that produce variability in intrinsic motivation. CET, which is considered a sub-theory of self-determination theory, argues that interpersonal events and structures (e.g., rewards, communications, feedback) that conduce toward feelings of competence during action can enhance intrinsic motivation for that action because they allow satisfaction of the basic psychological need for competence"(Ryan & Deci, 2000). The core claim of this theory is that external conditions can affect intrinsic motivation both in positive and negative ways. CET studies the effect of external events on IM and enjoyment, also explains in which conditions we can expect a rise in IM and where we will be hindering it. Empirical studies based on this theory reveal many interesting details on relation of IM and environmental factors. For example when someone is highly intrinsically motivated in a task, external incentives such as rewards can have negative effect on IM level. However if the IM is not high already, external rewards does not disturb IM, and might actually be helpful to improve performance.

Unexpected rewards yet did not have the negative effect even in high IM condition. Later same authors wrote a paper on "control and information" and its effect on IM. For example performance information can be presented so that user perceives it as pure information and feels that she has the authority to act upon it(or not). Same information can also be presented in a controlling way and make the user feel the pressure to improve or act. And of course, motivational and well-being outcomes of each condition differs greatly. Another interesting observation was that competence supporting elements are optimally effective when accompanied with feeling of autonomy. The reverse relation was also observed. Autonomy support was motivating when user felt the competence to make the right choice. An example of such conditions is when user feels overwhelmed by amount of choices she has.

2.2.3 Organismic Integration Theory

This mini-theory concentrates on extrinsic motivation. It seems that in adult life, extrinsic motivation is much more relevant and practical than intrinsic motivation. Because as adults we neither can love all the tasks and responsibilities we have, nor we can only do what we enjoy and feel intrinsically motivated towards. It is necessary to clarify that any sort of reasoning behind an action, already puts the motivation in place on the extrinsic side. Nevertheless motivation for a given goal or reason can be extrinsic but totally autonomous so that actor feels ownership of her action and decision and embraces the personal value and significance of the behavior . Perhaps having a good reason to perform an action is not enough to fully internalize it, but relating to that reason in a more personal level is important: "a smoker who

understands the 35 health benefits of cessation and wants to quit so that she might live to see her grandchildren grow up would display integrated regulation"(Vansteenkiste, Niemiec, & Soenens, 2010)

"Whereas the distinction between IM and EM was central to CET, this paradigm has been replaced by a distinction between autonomous motivation and controlled motivation. Autonomous motivation involves the regulation of behavior with the experiences of volition, psychological freedom, and reflective self-endorsement; the behavior has an internal perceived locus of causality. Both identified regulation and integrated regulation, in addition to IM, are autonomous forms of motivation."(Vansteenkiste et al., 2010)

Perhaps the most important contribution of this theory is the idea of organismic integration. OIT defines several levels of motivation(extrinsic) between intrinsic motivation on one end and amotivation on the other end. Table 2.1 explains each level in detail.

Behavior	Non self-Determined	...				Self-Determined
Motivation	Amotivation	Extrinsic Motivation				Intrinsic Motivation
Regulatory Style	Non-regulation	External Regulation	Introjected Regulation	Identified Regulation	Integrated Regulation	Intrinsic Regulation
Perceived Locus of Causality	Impersonal	External	Somewhat External	Somewhat Internal	Internal	Internal
Relevant Regulatory Processes	Non-intentional, Non-valuing, Incompetence, Lack of Control	Compliance, External Rewards and Punishments	Self-control, Ego-involvement, Internal Rewards and Punishments	Personal Importance, Conscious Valuing	Congruence, Awareness, Synthesis with self	Interest, Enjoyment, Satisfaction

Table 2.1: A taxonomy of human motivation (Ryan & Deci, 2000)

On the top of this table types of the motivation are listed: amotivation in the far left end, extrinsic motivation in the middle and intrinsic motivation at

the far right side. It also shows the gradual changes on the motivation moving from amotivation to intrinsic motivation. Second row show the regulatory styles related to each step of this gradual movement. Next row discusses the perceived locus of causality which means what is the source of the action in person's mind. Authors have distinguished levels of motivation in terms of their regulatory style and their causality.

Facilitating internalization is a big part of practical implications of this sub-theory. They argue that humans have natural tendency to internalize values and goals; and given the right conditions they will organically integrate them. This theory gives valuable directions on how to help this natural tendency to act better and faster.

2.2.4 Causality Orientations Theory

This theory defines a kind of surface level personality orientation towards autonomy and control. These personality trades are defined independent of any specific event or situation and are somehow constant to the person. Basic idea is that personality trades affects the outcomes of those external elements that was mentioned as influencing the motivation by previous sub-theories. Therefore one condition or event can be experienced differently by various people and affect their motivation differently. Deci, E. L., & Ryan, R. M. (1985) defined three groups of personality orientations(in relation to autonomy perception):

- **Autonomy oriented:** These people experience high degree of choice and interpret an event or situation more as autonomy supportive. They

prefer jobs and situations that involves more freedom, choice and greater initiatives from their side.

- **Control oriented:** People high in this orientation tend to shape their behavior as response to a control. This control can be personal and from inside or external. External events such as reward or punishments are important for them for example they may rely on deadlines or external pressure to motivate themselves or choose a path based on external incentives. The reaction to the controls can even take the form of rebellion, however in both cases responses are because of and based on the controls and therefore control oriented.

- **Impersonal oriented:** People with this orientation feel inability to perform tasks. They also feel incompetence and think that behaviours and outcomes are out of their intentional control.

2.2.5 Basic Needs Theory

Basic Needs Theory(BNT)(Deci & Ryan, 2002) as well as other sub-theories of SDT emphasizes on three basic needs, however BNT goes deeper on each of them and their relationships. It also specifies three dimensions of the social environment that support (rather than thwart) those needs: "autonomy-supportive (rather than controlling) contexts support autonomy, well-structured (rather than chaotic and demeaning) contexts support competence, and warm and responsive (rather than cold and neglectful) contexts support relatedness "(Vansteenkiste et al., 2010). It seems that meaningful choice is relevant in relation to autonomy but some studies showed that even

in non autonomous conditions having a meaning behind the task is helpful for internalization. The reason why a well structured environment is autonomy supportive is the same case of too many option being overwhelming. Well-structured environments are also easier to learn and therefore support competence. BNT also claims that three psychological needs are innate and universal, as opposed to some other theories that say that they depend on the upbringing or personality.(Vansteenkiste et al., 2010)

2.2.6 Goal Content Theory

This theory studies goals and their origins. It divides them to intrinsic and extrinsic goals. Intrinsic goals are the ones with personal and internal reasons such as personal growth and healthy lifestyle when these goals are not pursued because of an external motive. This categorization is important because studies show that intrinsic goals are more likely to satisfy one's basic needs and therefore have positive outcomes for the individual.(Vansteenkiste et al., 2010)

2.3 SDT in Digital World

It is important to have an overview of presence and use-cases of SDT in digital applications in general to see what can be taken and re-used and what is missing. Self-determination theory in digital world, mostly has been used in three fields of video games, gamification and educational tools. However involvement of SDT in each of these fields is different than other. For example SDT since its development, has been used in educational settings such as

classrooms. Therefore when it appears in educational digital tools, it comes as the transition of physical classrooms to digital environments. This view does not necessarily address needs of digital design, as much as it focuses on educational aspect. Nevertheless looking into these practices is informative and useful.

Moving to game, historical review shows that video game industry somewhat always followed STD's idea of motivation: they first started with competence, by introducing performance based goals. Old school points and levels were part of first game concepts. Then came the era of games with more options to choose, different goals to pursue and more dimensions to move in, which satisfied autonomy need. And last was the time of social games and community based gaming, addressing relatedness.

Although SDT has been studied in video games, most of those studies are psychological works aimed to show that SDT explains the video game interest. Their primary focus is not informing or evaluating design.

(Ryan et al., 2006) studies autonomy and competence in game-play. They show that perceived autonomy and competence in a game-play are associated with game enjoyment, further play preferences, and changes in well-being before and after play. This study uses different available games in order to simulate different conditions of need support.

Study of Przybylski et al.(2010) shows that cognitive evaluation theory (CET) predicts increased levels of game enjoyment, immersion, and preference for future play. Satisfaction of three needs of competence, autonomy and relatedness provides more general explanation to game play and game enjoyment than other trades such as aggression. (Przybylski, Rigby, & Ryan,

2010).

In video games relatedness is a well-used feature. But, because it is difficult to manipulate in lab conditions, it is much less studied compared to two other needs. Playing in teams or against others, shared effort for short or long term goals with real or virtual people can greatly improve the game-play experience. Tamborini and colleagues(2010) in a lab experiment on playing alone versus with a partner showed that playing with another player leads to more enjoyment (Tamborini et al., 2010).

It seems that design aspect of the games only comes in when talking about serious games or games with a purpose , games that although entertaining their main purpose is matters other than fun, such as training. This sort of games are studied in health related behaviour change applications. Thompson et al. (2008) explain how they designed a game to prevent diabetes using behavioural science. In this study they use SDT to inform some design elements such as feedback and a personalized goal setting(Thompson et al., 2008).

Exergames(games for exercise) as part of game with a purpose are studied in detail to see how and what design elements can support need satisfaction. Peng et al.(2012) conducted an experiment on their at home developed exergame and studied autonomy and competence supportive/non-supportive conditions. This experiment is one of the few that directly manipulated the game features mediating basic needs of SDT. These manipulations try to answer how the game features impact players' need satisfaction and game experience, while keeping all other conditions the same. Their choice of mediator for autonomy was possibility of customizing avatar and freedom in using

points for different weapons/powers. Competence support was implemented by a dynamic difficulty mechanism that would match player's performance , performance meter that would show how well player is performing and last feature was badges and collectibles. Their study first validated that these features actually affected players' perceived need satisfaction. They also showed a positive correlation between need satisfaction and enjoyment, future use interest, recommendation to others and rating given for the game(Peng, Lin, Pfeiffer, & Winn, 2012).

Overall studies of SDT in video games, although not numerous, more or less cover both holistic view of application use and studies focused on interactions and elements. Gamification on the other hand has been always accused of being too narrow and limiting. This criticism on gamification is on overall state of this field and not about its application of SDT. In fact SDT is brought to gamification to save the situation with its narrow application and limited reach. In the theoretical front gamification researchers have found out that in order to thrive as a field, gamification should be useful, interesting and engaging therefore, can not be limited to a set of predefined game elements, such as badges and leader-boards. Therefore they started a quest to find a theoretical base for it, to open the opportunity of finding new instants of gamification practices and/or elements instead of the old fixed and over-used set. Studies on negative effects of external motivators and the shift towards theoretical frameworks, soon suggested a change in focus of gamification practices from extrinsic motivation to intrinsic ones in research. Focusing on intrinsic motivation suggests a deeper user research to find out user's goals and motivations. General hype of user-centered design in HCI

also helped to promote user centered gamification practices in recent years(as an example see (Deterding, 2014)).

In (Deterding, 2012) Deterding tries to point out motivation as a goal for gamification. Later in (Deterding, 2013) and (Deterding, 2015) he defines a framework for gameful design based on SDT(more specifically CET) and idea of design lenses. This method tries to bridge the gap between theory and the design practice together. It is also not prescriptive or limiting and makes good connection between different steps of the design process(Deterding, 2015).

Recent survey(Seaborn & Fels 2015) on gamification showed a wide gap between theoretical frameworks of gamification and its applied studies and design practices. The theories and frameworks are not tested, further developed and validated by empirical studies, and at the same time applied research and practical designs do not follow the existing frameworks or rely on related theories. Although there were published studies that measured and reported effects of gamification, In my queries I did not find any reported study that designs or evaluates a gamified system based on SDT. Most of these papers measured performance outcomes(or behavioral outcomes) and some reported psychological outcomes as well(for more on effects of gamification see (Hamari, Koivisto, & Sarsa, 2014)). For example, empirical work of Mekler et al. (2013) studied effect of gamifying an image tagging tool. They used "intrinsic motivation inventory(IMI)" to measure psychological effects, but they did not mention any theory as base of their work(Mekler, Brühlmann, Opwis, & Tuch, 2013). My conclusion on gamification research is that, empirical studies of SDT in gamification if not totally non-existent , are not well addressed.

Chapter 3

Method

3.1 Study Design

The domain of the test tool/application did not play a major role in the goal and outcome of this study. Therefore selection process was based on basic requirements of the practical work and open to any category from work, education, entertainment etc. To choose the applications and the domain these criteria were considered:

- Applications must be freely available to ensure that this work is repeatable for other researchers. This would also minimize the complexity and cost of the implementation.
- Applications should not directly relate to/manipulate the motivation/ enjoyment/ happiness or any other trades under measurement.
- Doing a meaningful chunk of activity should be possible in the duration of the study (15-20 mins) so that user is able to go through at least one cycle of action-feedback and explore most of the system's features.

- System must have some minimum interactions and can not be overly passive (calming music, ..)

- Because of the study design and limitations in usage time and condition, system should not need other activities in parallel in order to make sense (example: excessive and workout practices, same time meditation practices)

Based on these considerations, I chose to explore applications in the domain of language learning as part of self improvement theme. A Google play inquiry was made among free applications with the "learn languages" search term. The top 1st and 10th applications were originally selected (see in appendix C, image C.1) and after investigating on the basis of above criteria, were proven to satisfy all of them. Both of the applications have PEGI3 label[1], which means it is suitable for all age groups.

3.2 Tools and Equipments

3.2.1 Devices

Tablet: A lenovo tablet(IdeaTab s6000-H) running android 4.2.2 was used to run the applications.

Headphone: A regular headphone(Philips SBC HLI45) was provided so that user could listen to audio(if any) without feeling uncomfortable. User had the option of using her own headphone or no headphone at all as well.

[1] `http://www.pegi.info/en/index/id/33/`

3.2.2 Application no.1: Duolingo

The top search result in google play was an application called Duolingo. Duolingo [2] is a language learning app that supports 54 different language courses across 23 languages. The app is available on iOS, Android and Windows 8 and 10 platforms with over 120 million registered users across the world. They also have a website for web access. In google play at the time, it had rating of 4.7 out of 5. In Appendix C you can find images of the application's screens.

3.2.3 Application no.2: Phrasebook

Phrasebook[3] was the 10th most popular application that came up with "learn languages" search term. It has rating of 4.4 out of 5 in google play. It is an application from Bravolol group and it is also available for iOS. This App has 13 languages and any language is accessible from any other 12 languages.

3.3 Experiment

Lab study was conducted in IDlab room A302 in Tallinn University. The room has a quiet atmosphere and minimal furniture (see figure E.1 in appendix E). During the study only the participant and myself as moderator were present in the room.

People actively learning a new language ,or interested in learning languages were invited to participate in the study. Also, students of Tallinn

[2] `http://www.duolingo.com`
[3] `http://bravolol.com/phrasebook/`

University's language courses were approached. The reason for this selective approach was to get closer to the natural user base and conditions that such applications might be used. Altogether 16 participants completed the test. They were asked to try out 2 applications in random order. Duration of the app-usage was estimated based on pilot tests to ensure that user has the chance of completing at least a chunk of meaningful interaction and trying out most of the apps options. The time was set to 8 minutes for each application, however, it was the suggested duration and user could use the app longer if she wished to do so. No specific tasks were given but they were encouraged to click around and explore the app. After each session apps would be reset to starting point and all the user data would be removed. User was asked to answer a set of questions in a browser window on the same tablet. Questions were a mixture of 8 general background questions, 4 items of "interest-enjoyment" subscale from "Intrinsic Motivation Inventory"(IMI) , 9 questions adapted from "player's Experience of Need Satisfaction"(PENS) ,11 items of "System Usability Scale"(SUS) and finally their future use preference and if they will recommend the application to others.(detail of scales will be in the scale section and the questions can be found in appendix B.

3.3.1 Procedure

- Greeting the user and explaining the the steps of experiment and what the goal is:

  "My study is about digital design. You will basically try out two applications and answer some questions after using each of them. Both

of the application are related to language learning and are freely available in google play. The whole experiment will take about 30 minutes. Tablet's screen or your in-app activities won't be monitored or recorded and your data will be cleared after you finish. You can ask questions any time and I will be here if you need assistance with anything. Here is the first application. It is called [the app name] And here is a headphone you can use to listen to audio. (will open the first application) It has several language options you can pick one as you wish and and may or may not follow a path. It is important to explore the App and try out different options. Now I will give you some time to use it, and I will let you know when it is time to finish"

- User plays with application for 8 minutes. The reason for not telling the user exact duration is to prevent them from looking at clock. They were informed that the whole duration of the experiment will be around 30 minutes.

- Informing user that the time is up and asking her to answer first set of questions:

  "You can stop any time now, but you can finish the task in hand if you wish to" "Here are questions [opening the page with questions]"

- And the same goes for the second application.

- Thanking the user and asking interview questions and engaging in a more detailed conversation based on their answers if necessary:

  "What did you like about [name of the app]"

"What you disliked/found annoying about [name of the app] or wish that was different"

- Thanking her again and answering any questions she might have.

Please refer to Appendix E for some photos from the setting.

3.4 Online Study

Other part to the experiment was an online study that has been launched in Duolingo community of users. Active users were approached directly from Duolingo discussion forum. Questionnaire was live for 3 days and had 154 responses. Questionnaire had a similar content as in-lab experiment with changes dictated by context: interview questions were added inside, system usability scale was removed due to length and the fact that regular users generally find a system usable. One question was added to determine if they use Duolingo on web or mobile/tablet. See Appendix D (images D.1 and D.2) for images of the questionnaire.

3.5 Scales

Interest-Enjoyment

This part was a subscale from Intrinsic motivation inventory(IMI)(Inventory, 1994). IMI is a tool to assess participants' subjective experience related to a target activity. Scale itself has 6 different subscales for measuring subjective interest/enjoyment, perceived competence, effort, value/usefulness, felt pressure and tension, and perceived choice. Each subscale could be used separately and adapted to the experiment. The enjoyment subscale contains 4 items of 7 point likert scale. IMI is validated and used in many studies as well as in game study of (Ryan et al., 2006).

Player's Experience of Need Satisfaction(PENS)

This scale is originally designed based on SDT for gaming context (Ryan et al., 2006). It claims to measure causal elements of an experience and predict fun, enjoyment and return of the user as well as popularity and ratings of given game. It has 3 parts for each variable of: competence, autonomy and relatedness. 3 items for each part with a 7 point likert scale.

System Usability Scale(SUS) (Brooke, J. ,1996)

This scale is designed to measure subjective usability of a system. It consists of 10 items in likert scale. SUS is a well used scale for various ICT systems. It is general and compact, and freely available as well.

Chapter 4

Results

4.1 General Statistics

Online study had 154 responses collected during 3 days. The target of this questionnaire were Duolingo users. The questionnaire was posted in official Duolingo community discussion page.

Gender: From this sum 83(53.5%) reported themselves as male, 66(43.2%) female and 5 (3.2%) other.

Age: 69 people out of 154 participant (44.5%) were under 20 years old, 44(29%) 20-30 years , 20(12.9%) 31-40 years and 21 (13.5%) over 40 years.

Education: 57(36.8%) people reported their education as high school, 40(25.8%) bachelor's , 28(18%) other, 22(14.2%master's) and 8 people's education(5.2%) was higher than master's.

Usage report:

110(71%) said that they use Duolingo everyday. 137(88.4%) said that they have tried other language learning tools. 95(61.3%) considered them-

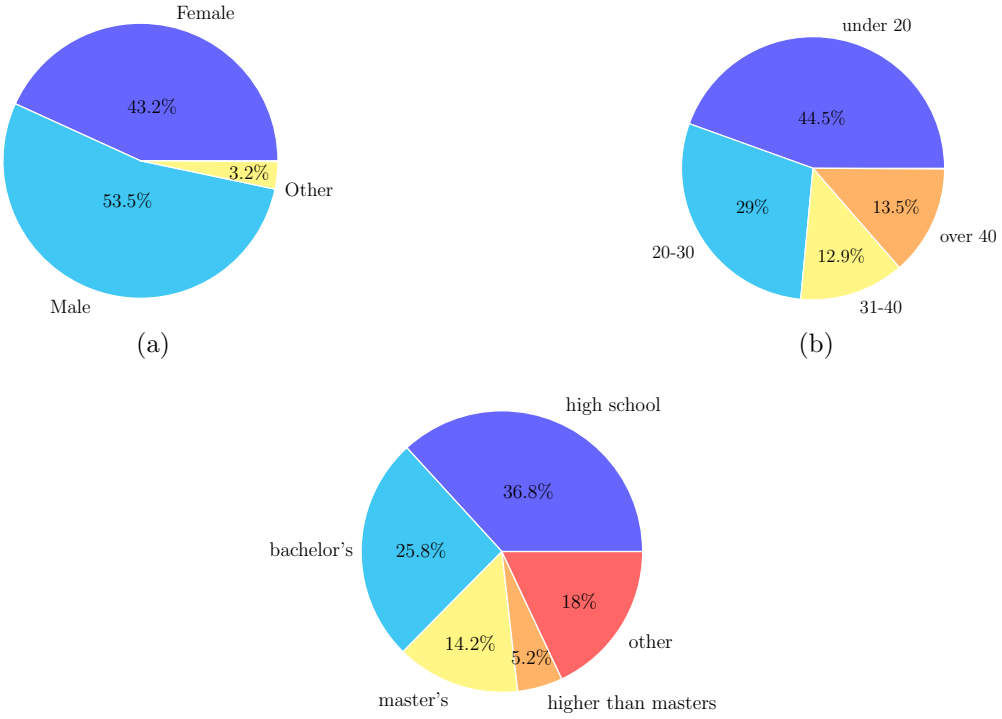

Figure 4.1: Visualisations of distribution of a)Gender b)Age and c)Education in online study

selves regular digital application user(computer, tablet, smart-phone applications)

The actual statistics of Duolingo users are not known. A similar survey[1] in 2015 was held in user community discussion page. Respondents to the survey were 71% male, 27% female and 2% other. Age Groups were reported as 39% under 20, 48% between 20 and 30, 7% 30-40 and the rest over 40. In that survey no information about user's education and technology use habits was recorded.

Although percentages(mostly in gender) differ in my survey and the older

[1]https://www.duolingo.com/comment/3938897

survey, it does not seem very critical for this study. There were no significant differences in any of needs under study(competence, autonomy, relatedness) or outcomes(for example enjoyment) for two groups of female and male.

In lab experiment : 16 people completed the tasks but data from one person was not usable, therefore removed for the rest of the analysis.

Gender: 6 male and 9 female. Age: 12 were in 20-30 age range, 2 were 31-40 and one under 20 years old.

Education: 6 people reported their education as bachelor's, 6 master's , 2 high school and one higher education.

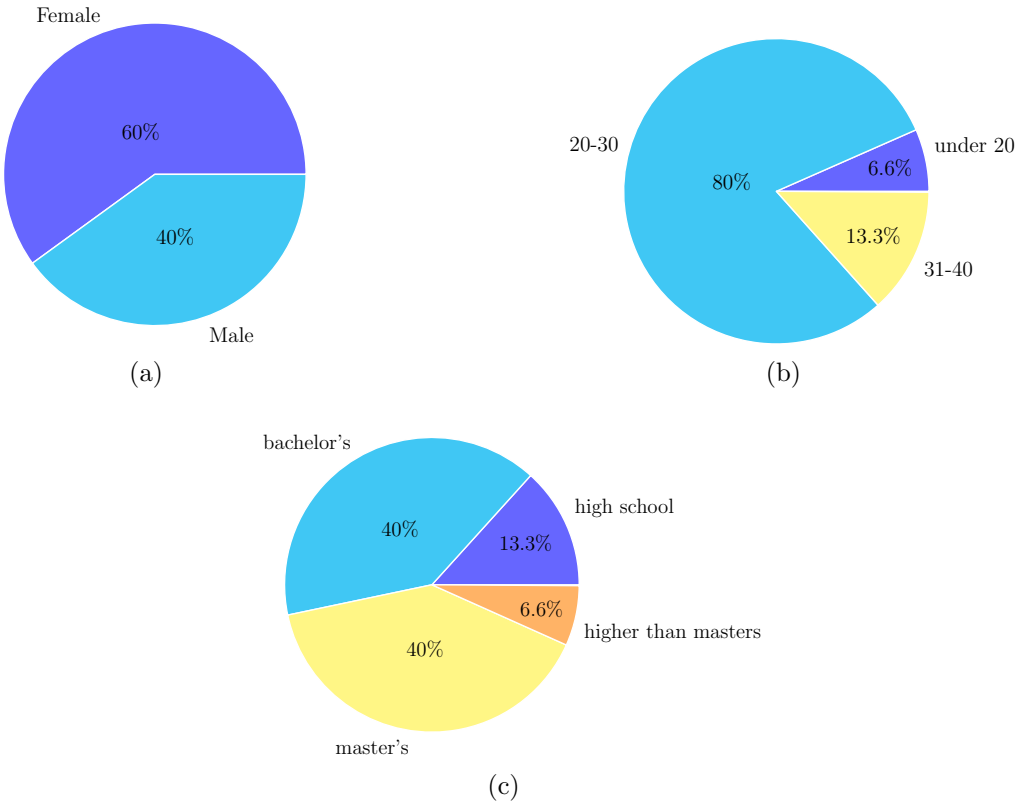

Figure 4.2: Visualisations of distribution of a)Gender b)Age and c)Education in lab study

Usage report: 12 out of 15 participants considered themselves as regular mobile phone/tablet app users. One participant was familiar with both application prior to the test, and 11 people were familiar with Duolingo but non of 15 participants were a regular user of any of two apps.

4.2 Validation

Enjoyment/Interest: this scale had Cronbach's alpha of 0.876 for online study and 0.966 for in-lab experiment.

PENS-Competence: this scale had Cronbach's alpha of 0.674 for online study and 0.84 for lab experiment.

PENS-Autonomy: this scale had Cronbach's alpha of 0.794 for online study and 0.74 for in-lab study.

PENS-Relatedness: this scale had Cronbach's alpha of 0.84 for online study.

SUS: Cronbach's alpha for SUS was 0.87 in lab experiment.

4.3 Hypotheses and Outcomes

Hypothesis 1 *User's perceived experience of need satisfaction in namely : competence, autonomy and relatedness, will predict three outcomes of : enjoyment , how often they use it(in online study) and willingness to recommend it to others and likelihood of using the apps in the future(lab experiment).*

A Spearman's rank-order correlation was run on the data to determine the relationship between perceived need satisfaction and mentioned outcomes. Below are the outcomes and related details.

Need satisfaction and enjoyment/Interest(online study):

Perceived competence and enjoyment/interest showed moderate, positive correlation with Spearman's correlation coefficient of 0.59 ($\rho < 0.01$).

Perceived autonomy and enjoyment/interest also showed moderate positive correlation with Spearman's correlation coefficient of 0.56 ($\rho < 0.01$).

Relatedness and enjoyment/interest showed weak correlation with Spearman's correlation coefficient 0.38 ($\rho < 0.01$).

Need satisfaction and enjoyment/interest(in-lab experiment):

Perceived competence and enjoyment/interest showed positive correlation with Spearman's correlation coefficient of 0.68 and significance level of =0.01

Perceived autonomy and enjoyment/interest also showed strong positive correlation with Spearman's correlation coefficient of 0.78 and significance level of =0.01

Need satisfaction and frequency of use(online study):

Since lab participants were not users of the apps, this part was omitted for lab experiment.

No statistically significant correlation between any of needs and frequency of use was found for online study.

Need satisfaction and likelihood of using the apps in future(lab experiment): Competence and likelihood of using the apps in future was strongly correlated with Spearman's correlation coefficient of 0.70($\rho < 0.01$).

Perceived autonomy and likelihood of using the apps in future showed moderate positive correlation. Spearman's correlation coefficient was 0.65 for this pair($\rho < 0.01$).

Need satisfaction and likelihood of recommending the App to others(online study):

Perceived competence and likelihood of recommending showed moderate, positive correlation with Spearman's correlation coefficient of 0.4 ($\rho < 0.01$).

Perceived autonomy and likelihood of recommending also showed weak positive correlation with Spearman's correlation coefficient of 0.27 ($\rho < 0.01$).

Relatedness and likelihood of recommendation showed very weak correlation with Spearman's correlation coefficient 0.2 ($\rho < 0.05$)

Need satisfaction and likelihood of recommending Apps to others(in-lab experiment):

Perceived competence and likelihood of recommending showed moderate, positive correlation with Spearman's correlation coefficient of 0.63 ($\rho < 0.01$).

Perceived autonomy and likelihood of recommending also showed moderate positive correlation with Spearman's correlation coefficient of 0.60 ($\rho < 0.01$).

Relatedness and likelihood of recommending was not possible to report since relatedness was missing in lab experiment.

Hypothesis 2 *There will be a significant difference between means of need satisfaction for three needs of : Competence, Autonomy and Relatedness, in two applications with difference in popularity. I expect the more popular*

	Competence	Autonomy	Relatedness
Enjoyment	0.59	0.56	0.37
Recommending	0.39	0.27	0.20
Frequency of use	No statistically significant correlation	No statistically significant correlation	No statistically significant correlation

Table 4.1: Correlations of need satisfaction for competence, autonomy and relatedness with enjoyment, likelihood of recommending and frequency of use in online study($\rho < 0.01$).

	Competence	Autonomy
Enjoyment	0.68	0.78
Recommending	0.63	0.60
likelihood of future use	0.70	0.65

Table 4.2: Correlations of need satisfaction for competence and autonomy with enjoyment and likelihood of recommending the App to others, in lab experiment($\rho < 0.01$).

application to have higher need satisfaction means.

Perceived competence in two applications: Paired t-test revealed that difference in mean competence for two applications is statistically significant in α =0.01($\rho < 0.01$). Mean competence for Duolingo was 5.68 and for Phrasebook 4.42 .

Perceived autonomy in two applications: Duolingo had a mean autonomy of 5.40 ,this number for phrasebook was 3.91 . Paired t-test showed that this 1.5 point difference in means is statistically significant.(α =0.01)($\rho < 0.01$)

Perceived relatedness in two applications: Relatedness was not possible to measure in the lab experiment since test subjects did not have enough time to build any relationships in either of applications.

Hypothesis 3 *Does usability correlate with need satisfaction for Competence and autonomy? I hypothesize that higher SUS value will predict higher*

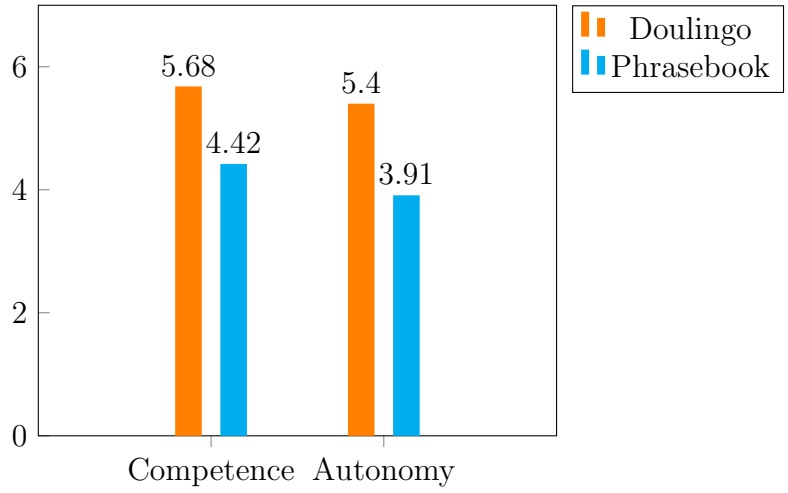

Figure 4.3: Avarage competence and autonomy for Duolingo and Phrasebook. The differences are statistically significant($\rho < 0.01$).

need satisfaction.

Need satisfaction and SUS: Spearman's correlation was ran for each pair to test this hypothesis.

Competence and SUS show a strong statistically significant correlation of 0.83 ($\alpha = 0.01$) in Duolingo in-lab tests. Same correlation test showed lower result of 0.72 for Phrasebook tests ($\alpha = 0.01$) .

Autonomy and SUS show no significant correlation in Duolingo in-lab tests. Phrasebook tests showed statistically significant correlation of 0.71 ($\alpha = 0.01$) between autonomy and SUS values($\rho < 0.01$).

Side question 1: Does usability differ for applications ranked one and ten?

Mean SUS usabilities, 68.5 for Phrasebook and 83.1 for Duolingo had statistically significant differences of 12 points in α =0.01($\rho < 0.01$). It is

in line with their popularity and rating. Duolingo in the rank 1 had better usability mean of 83.1 and Phrasebook in rank 10 had lower usability of 68.5.

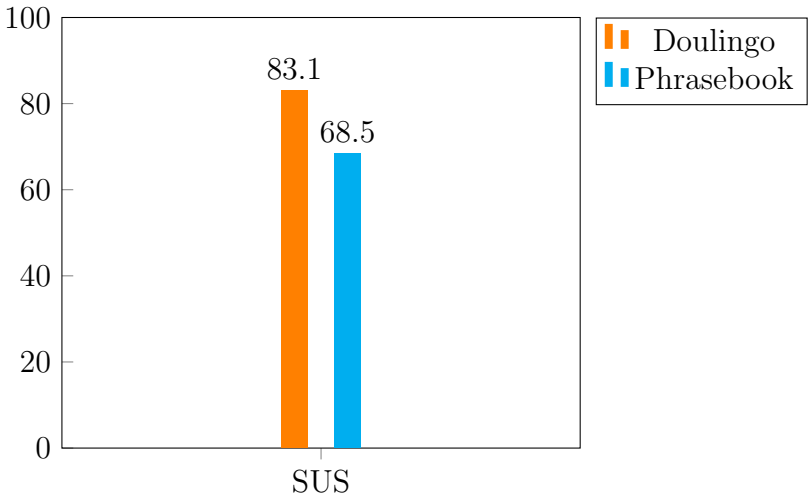

Figure 4.4: Comparison of system usability(SUS) means for Duolingo and Phrasebook(lab experiment)

Chapter 5

Discussion

5.1 Discussing Hypotheses and Results

To address three hypotheses (described in section 1.2) and two questions proposed in this work, two separate studies using two popular mobile applications were implemented. Both of the mobile applications are from the domain of language learning and they are freely available on *Google Play*. *Duolingo* was the application in the first rank with "learning languages" search term, and *Phrasebook* was the 10th most popular app in that search results. Online study targeted active Duolingo users and Lab experiment included people who were not active users of either of Duolingo or Phrasebook.

Hypothesis one claims that user's perceived satisfaction of competence, autonomy and relatedness needs will correlate with their enjoyment , how often they use the application(if relevant), if they will start using it in the future(if relevant) and willingness to recommend it to others. Results show that satisfaction of needs predict enjoyment in both lab and online study.

Lab experiment data show stronger correlations compared to online study. This might be due to the fact that lab study had more variation, while online study targeted only active users of Duolingo. Also online study was based on recalling the interaction while the lab experiment data was collected immediately after the use.

In the lab experiment people's preference for future use has strong correlations with both competence and autonomy. Among three needs, relatedness had the weakest correlation with enjoyment and likelihood of recommending Duolingo. It also has lowest mean compared to competence and autonomy(mean of relatedness,competence and autonomy Respectively 3.75 , 5.56 , 5.24) for Duolingo users. Relatedness not only in this work but also in previous studies proved to be a difficult trade to measure. Perhaps this feeling is harder to develop and it also needs longer period of interactions to be visible or even meaningful. Many studies excluded this need because of practical limitations. For example Peng et al. (2012) left relatedness out because of limited resources and the fact that it needed longitudinal study. It is possible that in a longitudinal study the changes in relatedness will be more visible.

However, it was curious that in the users' comments in open-ended questions(online study), features of relatedness(forums and comments, etc.) were the most mentioned features. This may indicate that the PENS scale does not effectively capture relatedness in this context. For example one item in the relatedness sub-scale is *"I find the relationships I form in Duolingo important"*; here *important* might be a strong word for describing a virtual relationship. Also people may not feel comfortable admitting to importance of such relationships. Although at this point this is only speculation, it is a

viable possibility that needs further exploration.

It is also worth pointing out that since the questions were answered in written form and there was no opportunity of clarifying the points, it is possible that people mentioned forums and comments in a different sense than relatedness.

In online study, need satisfaction and likelihood of recommending apps to others showed positive but weak correlations. Need satisfaction and frequency of use did not show any statistically significant correlations. Both likelihood of recommending and frequency of use had very high averages (respectively 6.20 and 6.42 out of 7 point scale) with very few low points. It seems that the respondent group was very uniform. Perhaps to capture differences in this case, different groups of users with more variance in their frequency of use should have been approached.

"Likelihood of recommendation" also could have been replaced with "if you have already recommended the app". This way more actual action of recommending could be captured.

Hypothesis two is about comparing two applications based on SDT. It expects more popular application to afford higher need satisfaction for its users, which proved true. Duolingo, the top ranked application had statistically significantly higher autonomy and competence means compared to Phrasebook(rank 10).

Google Play rank takes into account several measures such as number of installs and uninstalls, user ratings and download growth. And although

these results are based on one session use and do not say much about long term effects of them, the first encounters of user with the application is very important to ensure that in the future user will return to the application.

Hypothesis three looks into usability and its relation with need satisfaction for Competence and autonomy. Usability is defiantly a big part of user's experience with a tool. A good usability makes sure that user is able to pursue her goals in a fast and easy way. In the lab experiment usability showed strong positive correlations with autonomy and competence. Similar results were reported in previous studies for intuitive controls such as in (Ryan et al., 2006).

The perceived competence, in the applications subject of this study, can originate from three different sources:

- Usability: If the tool is hard to figure out, use, learn and recall; user will feel less competent.
- Interaction design: how a system implements competence boosting features such as feedbacks, impacts the perceived competence of the user.
- Content management: In this case if the learning material is too easy for the user because her language level is higher than what this tool provides; she will not feel boost of competence even if the system implements good interactions.

Based on this analogy, system usability is a big part of competence support and should not be neglected.

5.2 Interviews and Open-ended Questions

Research question 4: Can most liked and disliked features of the apps be explained by SDT?

At the end of each lab experiment session, participants were interviewed about what they liked and disliked about each of the two apps. Two simple questions were asked "what did you like about Duolingo/Phrasebook ?" "What you disliked/found annoying about Duolingo/Phrasebook or wished that was different" If the answer was unclear more follow up questions were asked to clarify the point. All the responses to these two questions were transcribed for both applications. Similar questions were asked in online study. In the analysis step I coded each block of response in three codes of "competence related", "autonomy related", "Relatedness related" and "other". Here a response block is the unit of speech that user is talking about one aspect of a single feature or has a specific focus. For example, people often answered those questions with more than one liked/disliked feature, which was made of several units.

There were several interesting repeating comments that did not necessarily fit into those three codes. I will mention those comments at the end, after need related comments, as they may also reveal some interesting points.

Features related to competence:

- Concept of *Duolingo being adaptive to user's level* came up several times during interviews. It is related to support for competence and providing the optimal challenge for the user. Duolingo would gradually

make the challenges harder if user successfully completes previous tasks, however, if she makes mistakes, system will repeat mixture of previous challenges to help the learning process.(lab experiment)

- For some of the absolute beginners though, the tasks were too difficult. Especially when user perceived the task as a "test" and not learning material. This might go back to user's personality trade such as orientations mentioned in Causality Orientations theory.(lab experiment)

- Another competence related comment was the placement test option in Duolingo. Users had the option to start as an absolute beginner or take a test to determine their level. Phreasebook not having a level label on the sections, or the fact that it is not sorted(easy to hard) was also mentioned as a minus.(lab experiment)

- Accepting all possible answers and being forgiving when user makes minor errors helped user to feel more competent. For example in the translation exercises there are more than one correct way of translating a sentence, therefore when a potentially correct variation is not accepted user felt frustrated. In the same type of exercise minor errors such as lower/upper case letters or minor misspellings were forgiven, in the sense that answer would be accepted but the mistakes were also pointed out. (lab experiment)

- Duolingo users mentioned seeing their friends' progress status as motivating.(however, I am not sure if this is competence related or it is more about relatedness). Some of these people directly mentioned competing

with their friends as a motivational force for them.

- Points, badges and levels were also mentioned as liked features in both lab and online study.

- *Streak* , the feature in Duolingo that shows how many days in a row user has reached her daily goals, was quite popular among active users. They found it motivating and engaging.

- Users of both groups of regular Duolingo users and lab participants found "immidiate feedback" an attractive feature.

- Many of Duolingo active users found "timed practice" where they have a limited time to complete tasks as fun feature.

Features related to autonomy:

- Freedom to choose what to study, freedom of browsing different topics and search-ability was seen with many as a positive feature of Phrasebook and on the opposite side users mentioned that Duolingo not letting them brows next topics before finishing the previous ones, was limiting. Both of these phenomena are related to autonomy support. However, the path that user is forced to follow in the case of Duolingo also closely impacts competence factor. The place you are in the path in Duolingo, is a sort of general progress indicator. The designers here have chosen to provide user with a linear path. Having only one possible path is in my opinion why people found limiting. Clear path and freedom although in many cases are somehow in competition, they

are not necessarily opposite. One solution could be providing multiple paths, or letting user design her path. This solution may still add some confusion(see next point), of course the extent of its benefit/harm should be studied in user tests. It is interesting to observe that much fewer people from Duolingo user community mentioned the freedom of choosing the course or topic as a problem.

- Clear path in Duolingo versus confusion of not knowing where to start in phrasebook was one of prominent comments. However, people admitted that the confusion only lasted for few seconds and they eventually found their way around. This can be an example of structure versus control as it discussed in many literature of educational application of SDT. As mentioned in previous points, not having a clear path designed for user, might affect competence support in the long run. Therefore this may not be the simple first encounter problem. In the case of Phrasebook user might miss the progress indicator role of the path as well as finding the system confusing at the first use. (lab experiment)

- Many users wish for being able to customize the exercises was mentioned several times. For example the option of changing proportions of excesses, if user feels that she needs more vocabulary exercise over listening etc.

Features related to relatedness:

- Comments section: for each sentence to translate there is a comments section that all users can discuss various aspects of the sentence from

its structure to fun facts. This feature second to forums, was one of most mentioned feature.(online study)

- Forums were mentioned more than any feature by Duolingo users. However, it was not always as liked feature. Nobody disliked forums itself , but many wished for more topical and organized discussions. So many different aspects of the forums were discussed in responses that I speculate forums may affect more than one need. (online study)

- Many users wished for chatting option with people who are learning same language or with native speakers.(online study)

- A few said that if they could pair up with others, that would motivate them in learning.(online study)

It is worth mentioning that most of the times when people talked about relatedness features above, they mentioned them being enjoyable and they did not necessarily mention building friendships with other or feeling belonging. In some cases they called forums and comments "helpful" and "useful".

Other repeating comments:

- Several of test users mentioned images and colours as an attraction point. Aesthetics seems to be a part of user engagement, at least in first few encounters with the tool. (lab experiment)

- Competing with friends in learning was one of the motivators users mentioned.

In general during lab experiment each user had at least one competence or autonomy tagged comment answering this question. In the online study

comments, forums and comments/discussions were mentioned often and competence related items were second most frequent.

Other interesting Observations

- Several people when mentioned collecting badges and points as their liked feature, felt that it was "Childish" or stupid. Many adults feel guilty for liking features that are fun and not as serious.

- The effect of personal orientation in their perception was visible in many cases. For example while majority perceived Duolingo questions as "repeat and practice", some users took it as "test" or evaluation. This view caused them to feel pressured to perform and less competent.

- It was quite interesting to observe that autonomy related negative remarks were much smaller portion of comments in online study compared to lab experiment. It seems that regular everyday users of the system somehow accept the limitations of the tool as characteristics of the system, while the first time users are more sensitive to them. During literature review, it was also quite interesting that in psychological experiments, autonomy support has been studied much more frequently than competence and relatedness. However, it is a much less represented need in digital world. It might be because manipulating autonomy in real world ,where in movements and options seems unlimited, is much easier and relatively straightforward. Specially considering that digital environments are inherently limiting when compared to real world. Of course this is changing due to advances in digital

technologies, and soon the situation might become reverse. This also shows the importance of studying this need in digital applications more thoroughly.

- *Immersion* was one of the popular features of the Duolingo for participants of online study. In Immersion people can contribute to the translation of a book or article by translating sentences, check others' translations and editing them. It has a wikipedia sort of structure where everyone can contribute as much as they can. You can see who was the last editor and you also will get notification when someone edits your translation.

5.3 Future Work

Current study had several limitations that can be improved in future with more thorough studies. Repeating the lab experiment with larger number of people in a longitudinal study can give better insight on long term use patterns and behavioural and psychological outcomes. Specially considering that SDT explains gradual change and internalization very well in its organismic integration theory. Behavioural outcomes such as performance was not target of this study therefore was not measured and analysed. However, behavioural outcomes can be important in many cases and should be studied in detail.

Online study in this work targeted a very specific group of people that were active users of the application. This was the result of limitation in accessing other user groups who were less engaged with the tool. Further work

including more variant user groups may reveal stronger and more interesting results.

Several new research worthy topics came out as a result of this work:

Relatedness as a psychological need should be studied further in digital applications. Effectiveness of PENS relatedness sub-scale may also benefit from some qualitative study.

The *Immersion* feature in Duolingo is attracting many users' attention. It would be informative to investigate why it is so interesting for users and what are the motivational affordances of it.

Studying features of digital tools in the light of their situated motivational afforfances (Deterding, 2011) may give us better understanding of their function and effects. Specially in larger and more complicated interaction units such as comments, that are affected from personality, environment and other external conditions.

References

Csikszentmihalyi, M. (1990). Flow: The psychology of optimal performance. *NY: Cambridge UniversityPress*.

Deci, E. L., & Ryan, R. (2002). Overview of self-determination theory: An organismic dialectical perspective. *Handbook of self-determination research*, 3–33.

Deterding, S. (2011). Situated motivational affordances of game elements: A conceptual model. In *Gamification: Using game design elements in non-gaming contexts, a workshop at chi.*

Deterding, S. (2012). Gamification: designing for motivation. *interactions*, *19*(4), 14–17.

Deterding, S. (2013). Skill atoms as design lenses for user-centered gameful design. In *Workshop papers chi2013.*

Deterding, S. (2014). Eudaimonic design, or: Six invitations to rethink gamification.

Deterding, S. (2015). The lens of intrinsic skill atoms: A method for gameful design. *Human–Computer Interaction*, *30*(3-4), 294–335.

Gagné, M., & Deci, E. L. (2005). Self-determination theory and work motivation. *Journal of Organizational behavior*, *26*(4), 331–362.

Graham, S., & Weiner, B. (1996). Theories and principles of motivation. *Handbook of educational psychology*, *4*, 63–84.

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work?–a literature review of empirical studies on gamification. In *System sciences (hicss), 2014 47th hawaii international conference on* (pp. 3025–3034).

Huotari, K., & Hamari, J. (2012). Defining gamification: a service marketing perspective. In *Proceeding of the 16th international academic mindtrek conference* (pp. 17–22).

Inventory, I. M. (1994). *Intrinsic motivation inventory (imi). the intrinsic motivation inventory, scale description.*

Kanat-Maymon, Y., Benjamin, M., Stavsky, A., Shoshani, A., & Roth, G. (2015). The role of basic need fulfillment in academic dishonesty: A self-determination theory perspective. *Contemporary Educational Psychology*, *43*, 1–9.

Mekler, E. D., Brühlmann, F., Opwis, K., & Tuch, A. N. (2013). Disassembling gamification: the effects of points and meaning on user motivation and performance. In *Chi'13 extended abstracts on human factors in computing systems* (pp. 1137–1142).

Peng, W., Lin, J.-H., Pfeiffer, K. A., & Winn, B. (2012). Need satisfaction supportive game features as motivational determinants: An experimental study of a self-determination theory guided exergame. *Media Psychology*, *15*(2), 175–196.

Peters, C., Castellano, G., & de Freitas, S. (2009). An exploration of user engagement in hci. In *Proceedings of the international workshop on affective-aware virtual agents and social robots* (p. 9).

Przybylski, A. K., Rigby, C. S., & Ryan, R. M. (2010). A motivational model of video game engagement. *Review of general psychology*, *14*(2), 154.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psy-*

chology, *25*(1), 54–67.

Ryan, R. M., Rigby, C. S., & Przybylski, A. (2006). The motivational pull of video games: A self-determination theory approach. *Motivation and emotion*, *30*(4), 344–360.

Skinner, B. (1971). Beyond freedom and dignity. new york: Alfred a knopf. *Alternatives to punishment*.

Skinner, E. A., & Belmont, M. J. (1993). Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year. *Journal of educational psychology*, *85*(4), 571.

Sutcliffe, A. (2009). Designing for user engagement: Aesthetic and attractive user interfaces. *Synthesis lectures on human-centered informatics*, *2*(1), 1–55.

Tamborini, R., Bowman, N. D., Eden, A., Grizzard, M., & Organ, A. (2010). Defining media enjoyment as the satisfaction of intrinsic needs. *Journal of communication*, *60*(4), 758–777.

Thompson, D., Baranowski, T., Buday, R., Baranowski, J., Thompson, V., Jago, R., & Griffith, M. J. (2008). Serious video games for health: how behavioral science guided the design of a game on diabetes and obesity. *Simulation & gaming*.

Vansteenkiste, M., Niemiec, C. P., & Soenens, B. (2010). The development of the five mini-theories of self-determination theory: An historical overview, emerging trends, and future directions. *Advances in motivation and achievement*, *16*, 105–166.

Chapter A

Eestikeelne kokkuvte (Summary in Estonian)

Digitaalsete vahendite laialdane kasutamine elu eri aspektides on muutnud motivatsiooni rolli digitaalsete lahenduste disainis üha olulisemaks.

Enesemääratlemise teooria on leidnud laialdast rakendust motivatsiooni uurimisel mitmetes muudes valdkondades, nagu näiteks arvutimngud.

Psühholoogiaalaste teooriate uurimise roll inimese ja arvuti interaktsioonis on toetada otsuste tegemist disainiprotsessides. Antud teooria rakendatavuse tõestamise korral inmese ja arvuti interaktsiooni kontekstis on vimalik vastavaid meetodeid kasutada disainiprotsessi ning hindamise eri etappides.

Käesoleva uurimustöö eesmrk on uurida enesemääratlemise teooriat kui võimalikku teoreetilist tausta endas motivatsiooni ja kaasahaaravust hõlmavate digitaalsete lahenduste disainis. Antud töö uurib enesemääratlemise teooria võimalikku rakendamist digitaalsete lahenduste disainis sarnaselt nagu seda on tehtud muudes valdkondades. Uurimustöös on kirjeldatud empiiriliste andmete väljaselgitamiseks läbi viidud laboriuuringu ning veebiküsitluse protsessi ning tulemusi. Mõlemad uuringud andsid lubavaid tulemusi, viidates enesemääratlemise teooriale kui motivatsiooni ja kaasahaaravuse lahtimõtestajale digitaalsete lahenduste disainis, mistõttu väärib see edasist uurimist.

Chapter B

Questionnaires

Questions related to PENS scale was removed in this view because of limitations in re-distribution to public.

Duolingo user questionnaire

Note: This questionnaire addresses people who are familiar with Duolingo!

Hello!
I am student of Human-computer interaction and this questionnaire is part of my master thesis study on design for motivation and engagement. It will help me a great deal if you fill this questionnaire. It should take about 3-4 minutes!
More :
-No personal, identifying information will be collected during the course of the study
-The data collected here will be anonymous and will be used only in the context of this study
-I aim to study duolingo's interaction design and not your performance or feelings, thus there is no right or wrong answer, please respond as freely and honestly as possible.

Demographics

please answer these questions about yourself.

1. **Your gender :**
   Mark only one oval.
   - ◯ Female
   - ◯ Male
   - ◯ other

2. **your age group :**
   Mark only one oval.
   - ◯ under 20
   - ◯ 20-30
   - ◯ 31-40
   - ◯ over 40

3. **your education level:**
   Mark only one oval.
   - ◯ high school
   - ◯ bachelor's
   - ◯ master's
   - ◯ higher
   - ◯ other

4. **how often you use Duolingo :**
Mark only one oval.

	1	2	3	4	5	6	7	
less than once per month	◯	◯	◯	◯	◯	◯	◯	every day

5. **Do you use Duolingo mostly on web or phone/tablet:**
Mark only one oval.

	1	2	3	4	5	6	7	
more on web	◯	◯	◯	◯	◯	◯	◯	more on phone/tablet

6. **How likely is that you recommend Duolingo to others(like: friends, family)**
Mark only one oval.

	1	2	3	4	5	6	7	
Not likely at all	◯	◯	◯	◯	◯	◯	◯	Very likely

7. **Have you tried any other language learning tools?**
Mark only one oval.

- ◯ Yes
- ◯ No

8. **you consider yourself a regular computer/digital tool/application user.**
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

Untitled Section

Based on your experience using Duolingo and activities inside it, for each of the following statements, please indicate how true it is for you, using the following scale:

9. **I enjoyed using Duolingo very much**
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

10. **Activities in Duolingo are fun to do.**
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

11. **I would describe activities in Duolingo as very interesting.**
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

12. **I think activities in Duolingo are quite enjoyable.**
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

PENS scale

Reflect on your experiences with using Duolingo and rate your agreement with the following statements:

13. [redacted]
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

14. [redacted]
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

15. [redacted]
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

16. [redacted]

Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

17. [redacted]

Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

18. [redacted]

Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

19. [redacted]

Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

20. I [redacted]

Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

21. [redacted]

Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

22. **Mark the options, you find most attractive in Duolingo**
Check all that apply.

- [ ] That you can see your progress, put your learnings in action and get immediate feedback
- [ ] The discussions part, comments, making friends and the community of Duolingo users
- [ ] Many learning options and paths that is provided by Duolingo, freedom to choose your path and learning pace
- [ ] Other: .....

23. **What is the main reason that you keep using Duolingo?**

.....

24. **What features you enjoy the most in Duolingo?**

.....

25. **If you had the power, what would you change in Duolingo?**

.....

v2 of app use questions

Hello!
-No personal, identifying information will be collected during the course of the study
-The data collected here will be anonymous and will be used only in the context of this study
-You may withdraw from this study at any time, for any reason.
-I aim to study the App and not your performance or feelings, thus there is no right or wrong answer, please respond as freely and honestly as possible.

* Required

Post-use questionnaire

Demographics

please answer these questions about yourself.

1. **Your gender :** *
   Mark only one oval.
   - ◯ Female
   - ◯ Male
   - ◯ other

2. **your age group :** *
   Mark only one oval.
   - ◯ under 20
   - ◯ 20-30
   - ◯ 31-40
   - ◯ over 40

3. **your education level:** *
   Mark only one oval.
   - ◯ high school
   - ◯ bachelor's
   - ◯ master's
   - ◯ higher
   - ◯ other

4. **you consider yourself a regular mobile App user** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

5. **Which app did you just use:**
Mark only one oval.

- ◯ Duolingo
- ◯ Phrasebook

6. **Were you familiar with this language learning tool before?**
Mark only one oval.

- ◯ Yes
- ◯ No

7. **how often you use this app :(leave this question out if you have not used this app before)**
Mark only one oval.

	1	2	3	4	5	6	7	
less than once per month	◯	◯	◯	◯	◯	◯	◯	every day

8. **Have you tried any other language learning tools before?** *
Mark only one oval.

- ◯ Yes
- ◯ No

Post-use questionnaire

Interest/Enjoyment subscale

Based on your experience using the app and activities inside the app, for each of the following statements, please indicate how true it is for you, using the following scale:

9. **I enjoyed using this app very much** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

10. **This activity was fun to do.** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

11. **I would describe this activity as very interesting.** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

12. **I thought this activity was quite enjoyable.** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

Post-use questionnaire

PENS scale

Reflect on your experiences while using the App and rate your agreement with the following statements:

13. [redacted] *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

14. [redacted] *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

15. [redacted]
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

16. [redacted] *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

17. [redacted]
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

18. [redacted]
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

19. [redacted] *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

20. [redacted]
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

21 [redacted] *

Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

Post-use questionnaire

SUS scale

Please read the statements and rate them based on your experience in the app you just tried. Try to input the first answer comes to your mind.

22. **I think that I would like to use this App frequently** *

Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

23. **I found the App unnecessarily complex** *

Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

24. **I thought the system was easy to use** *

Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

25. **I think that I would need the support of a technical person to be able to use this App** *

Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

26. **I found the various functions in this App were well integrated** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

27. **I thought there was too much inconsistency in this system** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

28. **I would imagine that most people would learn to use this App very quickly** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

29. **I found the system very difficult to use** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

30. **I felt very confident using the App** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

31. **I needed to learn a lot of things before I could get going with this system** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

Post-use questionnaire

Future use prefrences

please give your honest feedback

32. **I would like to continue using this app by myself in the future.** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

33. **How likely is that you recommend this App to others(like: friends, family)** *
Mark only one oval.

	1	2	3	4	5	6	7	
Not likely at all	◯	◯	◯	◯	◯	◯	◯	Very likely

34. **In general I would like to spend more time using this app.** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly disagree	◯	◯	◯	◯	◯	◯	◯	Strongly agree

Chapter C

Selection of Application

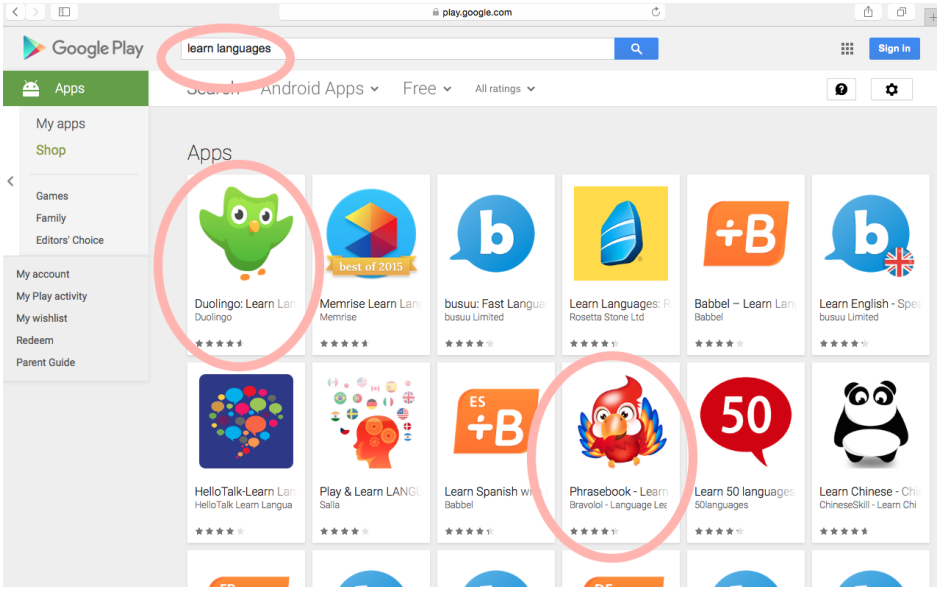

Figure C.1: Searching google play, with the search term "learn languages. First and 10th application in this list were selected for the study."

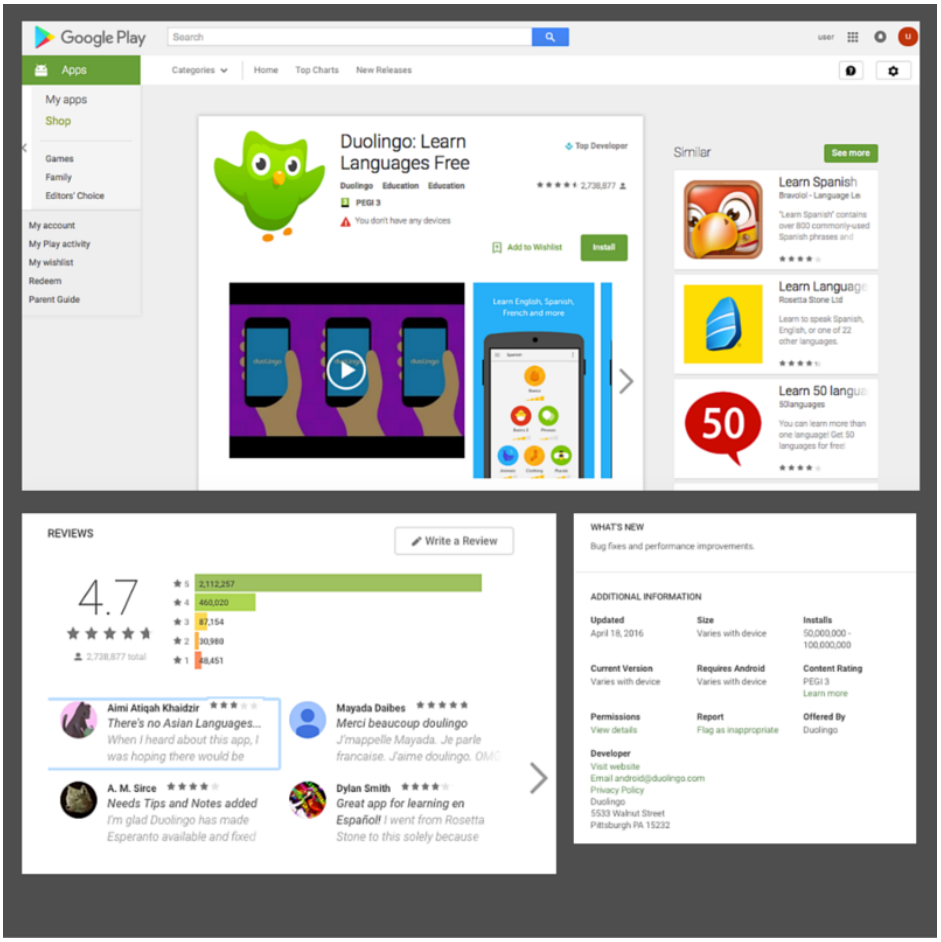

Figure C.2: The application Duolingo in google play.

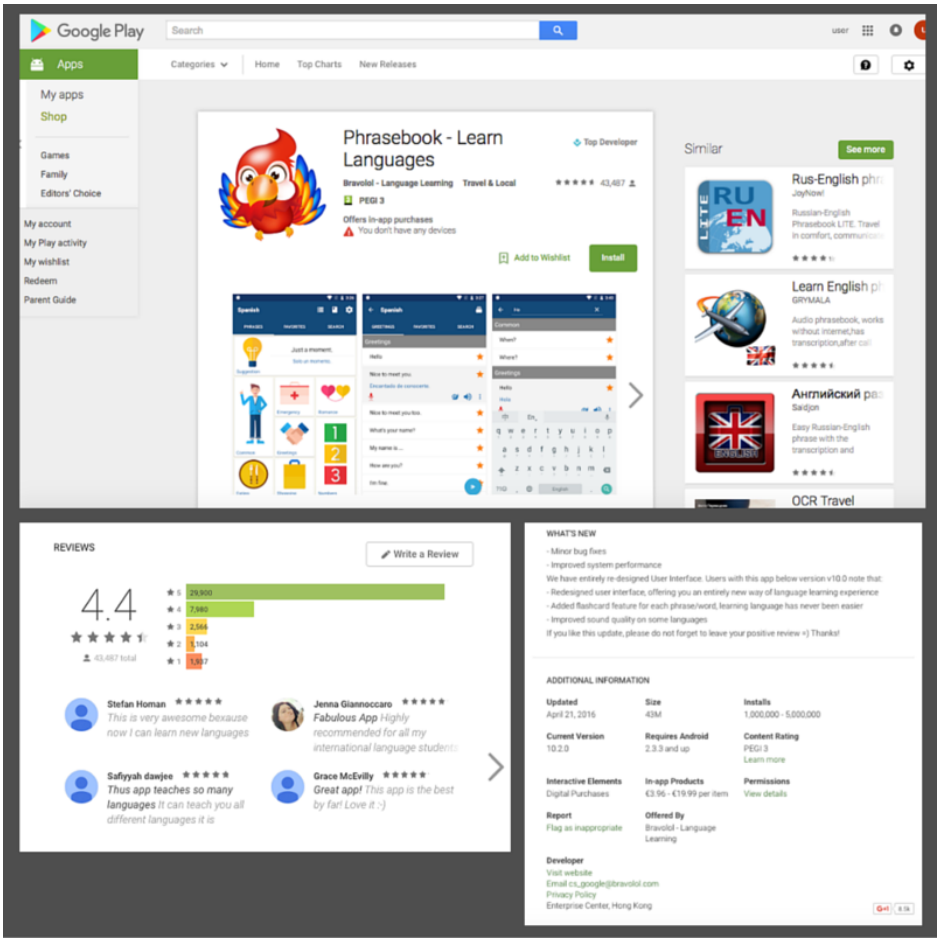

Figure C.3: The application Phrasebook in google play.

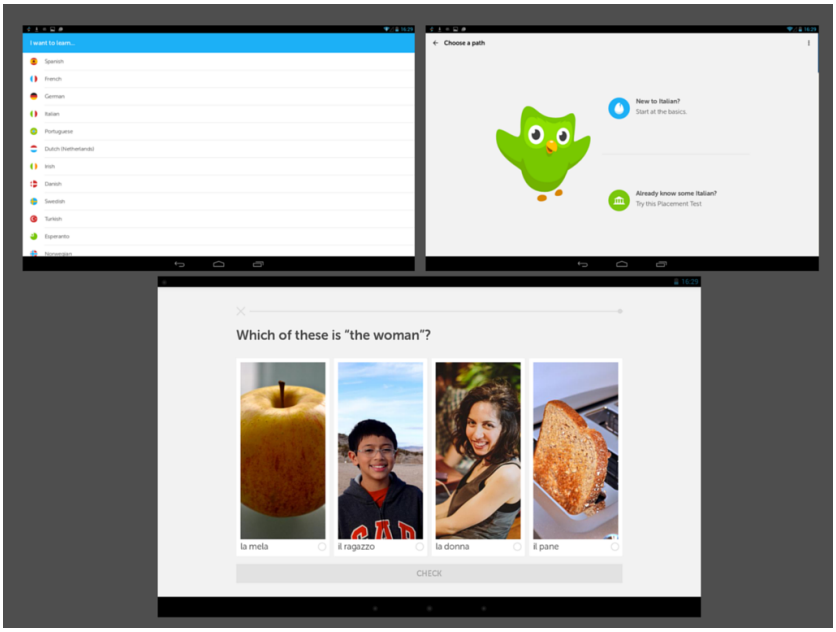

Figure C.4: In app screenshots of Duolingo on the test tablet.

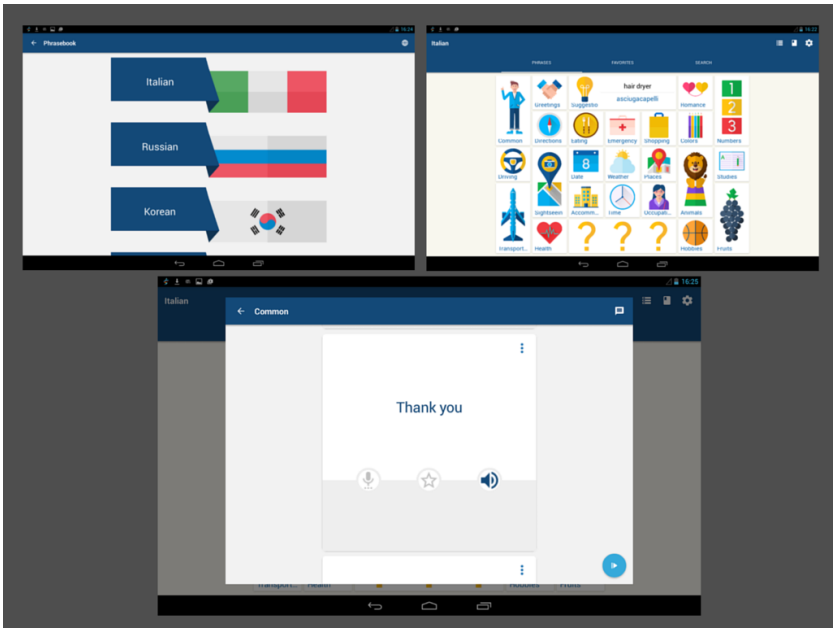

Figure C.5: In app screenshots of Phrasebook on the test tablet.

Chapter D

Questionnaire Screenshots

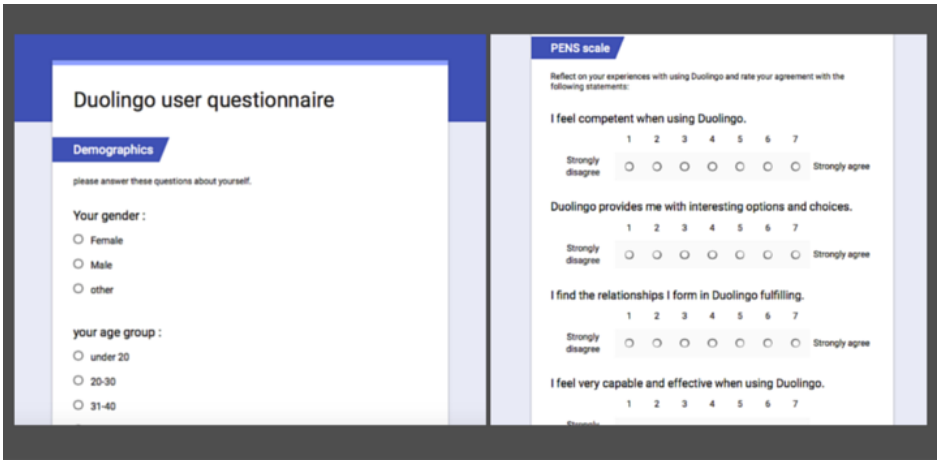

Figure D.1: Some screenshots of the online questionnaire, left: demographics and right: PENS scale.

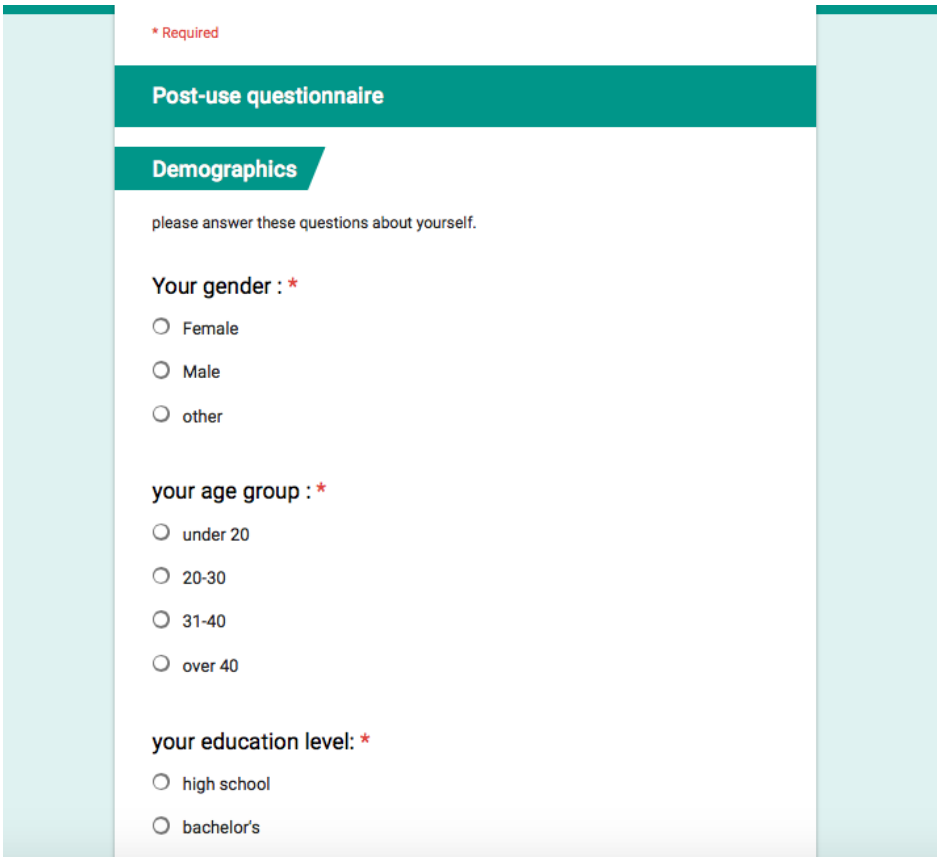

Figure D.2: Some screenshots of the lab study questionnaire, demographics.

Chapter E

Photos of Experiment Settings

Figure E.1: Room and conditions of the experiment. Third photo is a recreation of a user performing tasks.