

GAMIFICATION AND USER ENGAGEMENT IN SELF-LEARNING SOFTWARE

Juhani Tamminen

University of Tampere
School of information sciences
Information studies and
interactive media
Master's thesis
May 2015

UNIVERSITY OF TAMPERE, School of information sciences
Information studies and interactive media
TAMMINEN, JUHANI: Gamification and user engagement in self-learning software
Master's thesis, 60 pages, 3 appendix pages
May 2015

Gamification, or the use of game elements in non-game environments, is a field which has been growing rapidly in the recent years. The main focus of it is to engage users to activities or attach them to organisations, but often only to benefit the organisation, not the individual. User engagement and motivation on the other hand have been studied for example in the fields of management and psychology, where the focus has been to determine how feelings of engagement are formed.

The aim of this thesis is to understand how gamification affects user engagement and motivation as they are perceived by the users themselves. To achieve this a questionnaire study was conducted among the users of a gamified online language learning platform Duolingo. The results seem to indicate that even though in the case of Duolingo there are no organisational benefits to be gained from gamification, it is still implemented in a manner which does not induce feelings of engagement and can reduce intrinsic motivation. The results are reflected on recent gamification research and used to propose ideas on how gamification could be implemented better in systems such as Duolingo.

Keywords: gamification, learning, user engagement, motivation.

TABLE OF CONTENTS

1 INTRODUCTION.....	1
2 THEORETICAL FRAMEWORK.....	3
2.1 Games and play.....	3
2.2 Motivation, engagement and flow.....	6
2.2.1 Motivation.....	6
2.2.2 Engagement.....	8
2.2.3 Flow.....	10
3 GAMIFICATION RESEARCH.....	14
3.1 Background.....	14
3.2 Current research.....	17
4 METHOD AND TARGET GROUP OF STUDY.....	25
4.1 Duolingo.....	26
4.2 Selection of method.....	27
4.3 Formulating the questionnaire.....	30
4.4 Limitations.....	34
5 RESULTS.....	35
5.1 Psychological dimensions.....	36
5.2 Rewards and motivation.....	39
5.3 Flow.....	41
6 DISCUSSION.....	43
7 CONCLUSION.....	50
REFERENCES.....	52
APPENDICES.....	54

1 INTRODUCTION

The main concepts of this thesis are gamification, user engagement and motivation. Specifically there is an attempt to understand how gamification could be used to induce engagement and if there are other motivational aspects to consider.

Gamification in short means making tasks that are not inherently games appear more game-like, and thus make them more interesting (Deterding, Dixon, Khaled, & Nacke, 2011). In recent years gamification has risen from being a marginal phenomenon into a big business, which is being exploited to engage users in services and work (Deterding, 2014). The point of gamification has in many cases been to get users engaged in an activity that is otherwise seen as either boring or mundane. Organisations have used game-like elements such as points and rewards to get users hooked into using their services instead of competitors. Others have implemented score systems in workplaces so that employees would be more productive. In many cases gamification is only seen as something that would benefit the organisation that implements it, not so much as something that would be beneficial to the user (Bogost, 2011). On the other hand there are also those who claim that gamification could be used to benefit users and even improve quality of life, if for example workers can see their work as meaningful through gamification, instead of just working because they feel like they have to, or engage students in education to improve learning results. (McGonigal, 2011).

User engagement has been researched particularly from the viewpoint of engaging employees to their work. In this thesis the seminal work presented by Kahn (1990) is used as a starting point for understanding user engagement and defining how it links with gamification. The psychological dimensions of meaningfulness, safety and availability are considered to be important aspects in user engagement and they are compared to elements used in gamification in order to see which dimensions are in fact present. In addition to psychological dimensions, cognitive evaluation theory

is discussed to further understand how motivation is formed and what sort of effects rewards used in gamification have on it. Specifically the question of intrinsic motivation being replaced by extrinsic motivation as a result of rewards is looked at. (Deci, Koestner, & Ryan, 2001). Finally the concept of flow is used as an example of what could be called a high level of engagement, where users are fully involved in their task and almost immune to distractions (Csikszentmihalyi, 2013). Games are seen as something that induces the state of flow easily, so it stands to reason that gamification implementations aim to do that as well.

The target group of this thesis are the users of an online language learning platform Duolingo¹. It was selected because it is a good example of a gamified system where points and levels are added without connecting them to the learning task in a meaningful way. When a connection is not created users can perceive gamification only as an abstract layer of rewards separate from the task they are performing (Deterding, 2014; Nicholson, 2014). Duolingo is also simple enough to allow breaking it down to its core components and analysing them from the viewpoint of user engagement and motivation. There are also several other services that are similar to Duolingo, for example Lumosity, Codecademy and Steam,² all of which provide some form of point gamification without linking rewards to tasks. Seeing how the effects of gamification are perceived in one system should help understanding how they will work in other systems as well. User perspective is an important aspect of analysis, since a lot of gamification studies are made specifically from the viewpoint of organisations (Philippette, 2014). Understanding how users feel about gamification and its effects on their own engagement and motivation is an important factor in designing gamification that is engaging and beneficial to the user as well as to the organisation.

There are two primary questions and two additional questions this thesis is seeking to answer. First of all, 1.) does Duolingo as a service support the psychological dimensions as described by Kahn?

1 <http://www.duolingo.com/> (visited 5.5.2015).

2 <http://www.lumosity.com/>, <http://www.codecademy.com/>, <http://store.steampowered.com/> (visited 5.5.2015).

Secondly, 2.) is there a significant undermining effect for intrinsic motivation. Additionally, 3.) does Duolingo encourage long term learning and 4.) does using Duolingo induce a state of flow.

This thesis is structured so that chapter 2 provides a theoretical framework, first discussing definitions and play and games and then going through some seminal publications on user engagement, motivation and finally flow. Chapter 3 then presents background of gamification and some of the more recent research on the topic. Chapter 4 describes the target and method of study in more detail and chapter 5 presents the results. Finally Chapter 6 discusses the results, reflects them on the research presented in chapters 2 and 3 and presents some ideas on how gamification could be implemented in a more meaningful way in systems such as Duolingo.

2 THEORETICAL FRAMEWORK

Since games and play clearly relate to gamification, this chapter will first briefly discuss some of the attempts to define them, and what makes games different from other activities. Secondly, the concepts of user engagement, motivation and flow are looked into by examining some of the core theories in the field.

2.1 Games and play

Games themselves have been defined and redefined on numerous occasions and there are also a lot of attempts on trying to explain why games are played. The often cited early example of this can of course be found in *Homo Ludens* by Huizinga (1944), where it is stated that playfulness is pervasive not only in human behaviour, but in other species as well. So according to Huizinga, humankind has a built in need to play and frolic, in other words engage in activities that have no purpose other than the act of playing. These activities are present everywhere in culture and cannot be separated from it.

Huizinga's work was taken forward and expanded by Caillois (1958) with the addition of a more structured classification to the forms of play and games. Caillois presents a four-fold classification

which consists of *agôn*, or the players desire to win on her own merits; *alea*, the thrill that comes from games of chance; *ilinx*, the pursuit of vertigo or altered state of mind; and *mimicry*, referring to make believe and pretending to be someone or something else. There is also an axis ranging from *paidia* on one end to *ludus* on the other. Roughly interpreted the former refers to free play, like children playing cops and robbers where the rules are vague and subject to change, and the latter to a more structured form of play and games, like a game of football for example. Furthermore, Caillois presents six characteristics that he says an activity needs to have in order to be play:

1. Free, as in non-obligatory. Players can quit at any time.
2. Separate, the space and time of play need to be defined.
3. Uncertain, course of play cannot be foreseen. Player has a say in what goes on.
4. Unproductive, play cannot produce anything.
5. Rules, actions are judged only based upon the rules of play.
6. Make-believe, a different place in which the events occur as oppose to real life.

Suits (1978) presents additional definitions by building on what Huizinga and Caillois wrote before. He says: “Playing a game is the voluntary attempt to overcome unnecessary obstacles”, which indeed does summarise the essence of a game rather well. He also separates a *lusory* goal from a *prelusory* one in activities. The *prelusory* goal of a friendly football game for example is winning the game, but more importantly there is a *lusory* goal of playing and having a good time. Related to this, Suits also discusses play and professionalism, claimig that even if there is an ulterior motive to play, money and fame for example, it is only added on top of the underlying *lusory* attitude, without which the game would and could not be played.

Salen & Zimmerman (2004), in addition to discussing what the previously mentioned scholars had said about play and games, emphasised how users invest considerable amounts of time and effort in

both activities. Games have rules that need to be learned, specific settings or locations where players have to go and possibly specialized devices or equipment that need to be purchased. Time and money are spent on games seemingly for no reason.

Games are played voluntarily, and in many cases activities that would otherwise be simple are made difficult by arbitrary rules. This of course calls again for the lusory attitude, because otherwise there would be no point in playing any games. Rules can also be seen as something that makes things easier from a certain point of view. If a game is seen as something that has a beginning and an end, preferably with some sort of measurable outcome, that outcome can be easier to reach with the help of rules.

Time, effort and rules however are not enough to explain what makes games meaningful and interesting. Salen and Zimmerman (2004) suggest that games are made meaningful through interaction and decision making. Interactions between player and the game world must have some meaningful and clear effects in order to be interesting. Conflict within the game world gives the player problems to solve without the actual risk of harm that might otherwise be connected to situations that are commonplace especially in modern digital games. Secondly, decision making is seen as an important aspect in a game, particularly in a sense that the decisions need to have an actual and lasting effect on the game world.

Another crucial point is the difference between game or play and other activities. In this context the comparison can be limited to computer games and utility software, which can be crudely but efficiently divided in to their respective groups based on the users objectives. When a person uses for example a word processor at work, he has a clear goal he wishes to achieve, and the software is used as a tool to achieve it. In games the motivation comes from outside, usually from the game itself, and has no meaning to the user, or anyone, outside the game (Malone, 1982). In the light of the research presented it seems that playing games is inherently and fundamentally different from

work in its voluntary nature, freedom of choice and a set of rules that can be broken with no serious consequence.

2.2 Motivation, engagement and flow

This chapter presents some seminal publications on motivation, engagement and the concept of flow, which are all linked to play, games and gamification. Play and games provide their own motivation, engage users and can easily create a state of flow. Gamification on the other hand attempts to motivate and engage users in non-game activities using the elements of games.

2.2.1 Motivation

In a controversial paper published in 1971 Deci describes findings about the negative effects that rewards have on intrinsic motivation. Perhaps the most powerful critique to it was published in 1994 by Cameron and Pierce, which subsequently led to Deci, Koestner and Ryan publishing a response in 1999 and a revision of that in 2001. It is not in the scope of this paper to discuss further how the original and the critique differ from each other, but it can be said with some certainty that there is a sufficient scientific consensus supporting the original work to use it as a reference here.

Cognitive evaluation theory (CET) states is that intrinsic motivation stems from individual's need for competence and self-determination (Deci et al., 2001). All events that undermine those two are deemed to lessen intrinsic motivation, which is exactly what rewards seem to do. Deci et al. explain that as the “locus of causality” moves from internal towards external, the feeling of self-determination is lessened and vice versa. It is also worth noting that Deci makes a point of always referring to perceived competence and perceived self-determination, underlining that both are subjective experiences.

CET divides rewards to two groups, informational and controlling, and attempts to categorise them accordingly. Generally informational rewards are considered to increase self-determination and

controlling rewards to decrease it. Simply put, controlling rewards have the tendency to alter behaviour as people hope to get rewarded.

Rewards are also divided to verbal and tangible ones. Verbal rewards are also referred to as positive feedback in research literature. In short, verbal rewards are usually seen as informational, though in some cases searching for positive feedback may lead to altered behaviour. With verbal rewards, interpersonal relationships affect how they are received. Feedback from an authoritative figure for example is more controlling than that from a non-authoritative one. Tangible rewards on the other hand are primarily considered to be controlling, and are further divided to task-noncontingent, task-contingent and performance-contingent rewards (Ryan et al., 1983, as cited by Deci et al., 2001).

Task-noncontingent rewards require no participation as such in a task, and are rewarded for merely showing up, and according to Deci they are the exception to tangible rewards being controlling. Task-contingent rewards are again divided into engagement-contingent and completion-contingent. The former requires a person to work on the task, but nothing more than that. Therefore it is considered to have no positive effects in terms of competence or self-determination and is purely controlling the user. As Deci et al. put it: “there will be nothing to counteract the negative effects of control.” Completing a difficult task on the other hand implies competence. Hence, completion-contingent rewards can be considered to have some positive effect to counter the controlling aspects. Overall however, rewarding completion still directs user behaviour. Finally, performance-contingent tasks are even more controlling than completion-contingent ones, since they require a certain level of results to be achieved. However a high level of performance also implies competence, so there are again some aspects to counteract control. Overall control is still the primary aspect of these rewards as well. Illustration 1 roughly outlines the relationship of reward types from informational to controlling.

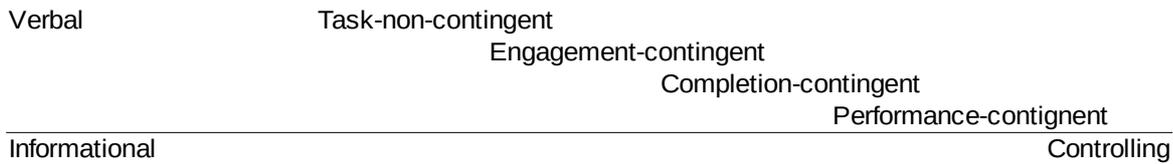


Illustration 1: Informational and controlling rewards.

As with verbal rewards, interpersonal relationships and the entire environment are again important in tangible rewards. If the environment is highly demanding and controlling as such, it leads to rewards being perceived as controlling as well. Supportive and error-allowing surrounding lessens the controlling affects and supports self-determination.

2.2.2 Engagement

The psychology of user engagement has been studied fairly thoroughly in various fields. Kahn's (1990) paper on engagement in the workplace is perhaps the most cited, and can be considered something of a cornerstone of study. It provided the first definition of engagement, though in this case the field of study was only limited to job engagement. The definition states that engagement means

The harnessing of organisation members' selves to their work roles; in engagement, people employ and express themselves physically, cognitively, and emotionally during role performances.

One of the key concepts in this definition is the idea of role. Kahn points out that people have different roles for different contexts, and that individuals invest their own personalities to these roles on varying levels (self-in-role). Different roles can be taken on throughout the day, and they change between tasks. The condition and context of tasks determine roles and the level of personal investment. In addition, roles that other people assume also affect each other, creating a network of communicating roles. The level of personal investment also dictates the level of personal engagement. The reason for assuming these roles is that working life is “ongoing, emotionally charged and psychologically complex.” Therefore people create roles that enable them to distance themselves from their work, allowing the opportunity to decide how much risk to their own persona

they are willing to accept. By calibrating their roles, people can also position themselves in a comfortable place within the community.

As the level of physical, cognitive and emotional expression within a role begins to raise, the “preferred self” emerges and the person becomes more engaged. The psychological dimensions of engagement are effort, involvement, flow, mindfulness and intrinsic motivation. They in turn express themselves through creativity, use of personal voice, emotional expression, authenticity, non-defensive communication, playfulness and ethical behaviour. It should be remembered that according to Kahn, even when a person is fully engaged in an activity, in other words invests herself fully in a role, it is still a role. This is an important realisation when evaluating and designing user engagement.

When talking about engagement, the other side of it should also be presented. In disengagement a person does not invest himself in a role, but rather keeps a distance. Compared to engagement, that leads to somewhat opposite psychological dimensions that can be automatic or robotic, burned out, apathetic or detached and effortless. Those again manifest in behaviour that is defensive, impersonal and emotionally inexpressive, bureaucratic, self-estranged or closed.

Kahn also conducted two studies to find out the reasons behind job engagement. He describes meaningfulness, safety and availability as the three psychological dimension that encourage people to invest themselves in their roles, and each of the three is further divided to subcategories which are briefly explained here.

Psychological meaningfulness can be described as the return of investment of self in role. One aspect of meaningfulness are task characteristics, which need to be challenging, allow some creativity and offer clear goals and some autonomy. Secondly, role characteristics are important in the sense that the assumed role needs to encourage engagement. If a person's role is very low in the workplace hierarchy, it is unlikely they will be highly engaged in their work. Lastly work

interactions such as personal interactions and feeling of appreciation within the group promote engagement.

The second dimension of psychological safety relates to the social setting of a workplace. Interactions need to be non-threatening and predictable, allowing individuals to invest themselves in them without the fear of getting hurt. Interpersonal relationships should allow trying something and failing without the danger of being reprimanded or ridiculed by others. Groups always are more than the mere sum of their parts, so group and intergroup dynamics come in to play in a sense that both formal and informal roles emerge within groups and again individuals position in the hierarchy determines the level of engagement. Finally organisational norms also matter with psychological safety, since there always are certain modes of operation that cannot be breached without consequence, and if those norms do not fit a person's preferred self, engagement is unlikely.

The last dimension of psychological availability is the sense of having physical, emotional and psychological resources available for engagement. Both physical and emotional energy are required for one to invest oneself in a role, and distractions can drain those energies. Simply put, tired employees are less likely to be engaged in their work. Insecurity at the workplace can have a great impact on work performance and engagement. Not only does it drain emotional energy, but it also creates a defence reaction where people can disengage from their work in order to avoid getting hurt. Excessive self-consciousness, as a result of being under continuous scrutiny for example, also leads to lessened likelihood of engagement. Finally, since work does not take place in a vacuum, outside life can affect engagement both in positive and negative. Positive events naturally increase person's energy levels, but negative events may cause preoccupation and distraction.

2.2.3 Flow

When the concept of flow was being formulated, there was a general notion that work is for the most part an inherently unpleasant activity (Csikszentmihalyi, 2013). In essence, people only work because they have to, and spend their leisure time with something that they actually enjoy. Hence,

giving rewards for carrying out dull activities was the only way of making those activities more interesting. However as Deci (2001) suggests, this may lead to the reward in fact becoming the sole purpose of doing a task. However as both Deci and Csikszentmihalyi point out, this method of stick and carrot is somewhat pervasive to society. For example children who manage to sit quietly in school are rewarded good grades for behaving in a manner that is rather unlike most children. And later on of course money is rewarded to those who show up at work and manage to perform in a satisfactory manner. According to Csikszentmihalyi, working only in hope of a reward can result in a deepening division of leisure and work and to a circle of needing more rewards. The more one complies to behaving as the rewards require, the less the activity itself is liked. That in turn leads to a need for increasing rewards.

In the core of Csikszentmihalyi's work there are questions of enjoyment and motivation. First of all, he asks how can intrinsic motivation be measured? Responses to external rewards can more or less easily be seen and measured, but intrinsic workings are more difficult to observe. Secondly, what is it that makes activities enjoyable? Sometimes the enjoyability may even seem contradictory, like for example when a long distance runner runs in spite of the pain and for seemingly no reason. Certain activities then seem to be enjoyable as such, without the need for any external rewards.

Csikszentmihalyi discusses autotelic people, activities, and experiences. In other words some individuals are more susceptible to the state of flow and similarly certain activities are more likely to induce it. An autotelic experience is the personal perception of the state of flow. He then lists eight reasons for enjoying autotelic activities, ranging from more intrinsic to extrinsic.

1. Enjoyment of the experience and use of skills.
2. The pattern, the action and the world an activity provides.
3. Development of personal skills.
4. Friendship and companionship.

5. Competition and measuring self against others.
6. Emotional release.
7. Prestige, regard and glamour.

Overall in his study it seems that the reasons deemed more intrinsic are also the more important and that unsurprisingly the perception of enjoyability is related to the task. Autotelic experiences are then divided to five groups based on what sort of activities they provide or require.

1. Friendship and relaxation.
2. Risk and chance.
3. Problem solving.
4. Competition.
5. Creativity.

What is perhaps the most important finding here so far is that activities that are novel, require problem solving and offer challenges are seen as more rewarding than activities that are for example previously known. It should also be pointed out that competition can be seen as a form of challenge, as well as exploring the unknown. Also breaking down social norms seems to have an effect that brings forth a feeling of closeness that would not otherwise be achieved. On the other hand, competition seems to lessen the feelings of friendship and relaxation, especially in so called zero-sum games, where there always has to be a loser in order for there to be a winner. At the same time however, such games when played in teams also encourage friendship among team members.

Generally it seems that the state of flow is often sought out for the sake of itself. Some activities are better at inducing it than others, games and play for example excel at that, but overall tasks that require some level of creativity or activities that cause a frenzy of some sort work well with

generating flow. In Caillois' (1958) classification of play and games, such frenzy is referred to as vertigo and is seen as an aspect of play.

The core elements of a flow experience consist of awareness and attention, clear goals and feedback, loss of self and not being afraid to lose. In games for example rules define all the aspects of a world that need to be taken into account, thus making it easier to focus on a task. On the other hand pursuing victory can work as a distraction from the actual task, as winning becomes the primary objective. This latter finding would seem to go together well with what Deci et al. (2001) suggested about controlling rewards altering behaviour and in this case becoming objectives on their own.

Having clear goals and receiving immediate feedback seem to go somewhat hand in hand with losing oneself in the activity. According to Csikszentmihalyi, a state of flow induces almost automatic responses to stimuli, which in turn would seem to go well together with the idea that when one loses the perception of self, one also loses a layer of cognitive processing. In this sense it could even be argued that a state of flow is the highest form of engagement, where users forget themselves and invest their preferred selves (Kahn, 1990) fully in the task at hand.

In the light of his research Csikszentmihalyi also suggests that the separation of work from fun is unnecessary or even harmful. The commonly held notion that something that is fun cannot be productive should be abandoned in favour of an approach where the importance, and thoroughness that enjoyment, of tasks is made plain to those performing it.

There are however some downsides to flow. It can for example be addictive in a way. Some participants in Csikszentmihalyi's study reported a need to reach flow that sometimes was actually hindering their performance. In other cases it was mentioned that reaching the top level, so to speak, of any task reduced the experience of flow, making it harder to concentrate on a task. Simply put, when a task becomes routine, it becomes boring and no longer offers challenges at the optimal skill

level and does not fulfil the requirements for flow. Inversely, a task that is too difficult and leads to many mistakes increases anxiety, hindering flow on the other end of the scale.

In the perhaps most common description of flow Csikszentmihalyi says that flow is induced when the challenge of a task is at an optimal level compared to individual's skills. What makes reaching flow difficult then is not understanding where it comes from, but rather how to compensate for the individual differences. Flow is a highly subjective state and as such, very difficult to predict. Skills and personality traits of a person may affect reaching flow more than design of the task.

3 GAMIFICATION RESEARCH

Gamification research has taken off, so to speak, in the past years. This chapter first presents the background of gamification and then some of the more recent research. A considerable portion of the chapter is based on the 2014 book *Rethinking Gamification*, edited by Fuchs, Fizek, Ruffino and Schrape.

3.1 Background

Defining what gamification is and is not seems to be an ongoing discussion that is, if anything, becoming increasingly polarised. Even the very definition of what gamification means seems to be disputed. The definition presented by Deterding et al. (2011) is in any case still a good starting point. They say that gamification is “the use of game design elements in non-game context.”

As was mentioned earlier, games are widely regarded as a part of culture. The cornerstones of studying games and play have been laid in *Homo Ludens* by Huizinga (1944) and *Man, Play and Games* by Caillois (1958), who for their own part have showed that play is not only universal among humans, but permeate almost throughout the entire animal kingdom. Games are played all over the world as a simple pass time or as a professional sport, either on tabletops, electronic devices or by running, riding or skiing in the real world. Even those who claim not to be interested in games often have to admit that they do at take part in the lottery at least every now and then.

In that sense it is difficult to understand why gamification is considered to be such a novel idea. And as Nicholson (2014) points out, people are in fact used to receiving abstract rewards for their actions in many fields of life. He claims that for example getting grades for school work is no different from getting points in a gamified system.

Handing out abstract rewards that have little or no meaning to people outside their own field is by no means a new idea. Medals and decorations are handed out for good performance in sports and in the military for example, and even the ancient Romans rewarded accomplished gladiators with medals (Köhne, Ewigleben, & Jackson, 2000). The American Boy Scout movement has a reward system based on badges that have very little meaning to outsiders. In addition, different reward systems for loyal customers (frequent flyer miles etc.) can be seen both as rewards for doing something and as a way of committing a person to a certain organisation, company or product. From the viewpoint of the users, they are rewarded for behaving in a certain way, and from the viewpoint of the organisation, they get loyal customers with very little effort. All the examples mentioned here can be seen as a form of gamification, a predecessor if you will, to the current development happening in the digital world. Schrape (2014) offers an explanation for this by saying that gamification is in fact a form of governing and controlling people to behave in a certain manner. Instead of a threat of punishment there is a reward system of point and badges. Positive feedback is accepted even if it means, as Schrape puts it: “[accepting] total surveillance”.

Usually when gamification is introduced to a normal workplace, it is driven by the benefits of the organisation, not the individual. As an early example of that kind of thinking Nelson (2012) presents precursors of gamification in the Soviet Union and later in the United States. According to Nelson, Lenin had proposed a system where individual workers, groups or entire factories would compete against each other. Based on those ideas a wide variety of reward schemes were implemented. Since the underlying idea was to replace capitalistic, money driven competition with something else, rewards were abstract things like points and medals. The later example from United

States has more to do with playfulness and so called funsultants, who were aiming to make the workplace more fun in the 1990's and early 2000's. Academic views on the subject are for the most part critical, saying that they were only implemented to control employees under the pretence of fun (Nelson, 2012).

Even though gamification has recently been discussed almost exclusively from the viewpoint of engaging customers, improving efficiency and altering behaviour, there are alternate views. McGonigal (2011) is one of the strongest advocates of gamification being something other than a tool for getting people to behave in a certain manner. She sees it as a tool for changing the world through harnessing our instinct to play. Contrary views have been brought forward by for example Bogost (2011), who says that in the current format gamification should in fact be called exploitationware. Systems where a simple game mechanic is added to a non-game environment should not, in his opinion, be linked to games in any way. Bogost instead talks about serious games, giving them credit as possible drivers for social and political change, as long as the games are designed as games and not as simple mechanical devices for collecting points.

There have indeed been some successful examples of implementing games where the actual gameplay is used to solve complex problems. For example a game called Foldit³ creates puzzle games that represent proteins. The player is asked to figure out how a particular protein folds. Best solutions are then tested in a laboratory to see whether or not they are viable. Another example of games with additional purpose are those where players are placed in a situation resembling the real world. The decisions made affect the world as they would in real life, and the point is in most cases to educate people about the effects of their actions⁴. This however does not fit to the definition of gamification, as presented by Deterding et al. (2011).

3 See <http://fold.it/portal/> (visited 30.10.2014).

4 See for example <http://benshapi.ro/projects/trails-forward/> (visited 30.10.2014).

What must be understood is that the current gamification systems for the most part presume games to be inherently enjoyable and work being boring. But using Caillois' (1958) classification of *paidia* and *ludus*, current gamification relies heavily in the structured forms of *ludus*. Such structures may work reasonably well in making unmotivating tasks more interesting, but to create a truly meaningful experiences the free play of *paidia* is required.

Hence, bringing game-like elements to a mundane task does not automatically make it interesting, nor is there a direct link between reward and motivation. Game design elements do in any case afford motivation. The concept of affordance, as explained by Norman (2002), refers to “the perceived or actual properties of the thing.” Games however are interesting and engage users for different reasons than what it seems gamified systems have done so far.

In the case of gamified systems the affordances they provide depend both on their design and previous interaction. They need to include challenges on an appropriate level for the user, but the user also needs to have some understanding about how gamified systems work. And most importantly, gamification needs to be designed from the bottom up to be meaningful to the user.

Game design then can be used in creating ludic experiences around tasks that are not inherently playful. But the practices of design need to be applied from the start, applying game mechanics, which leads to dynamic interaction, which in turn leads to a meaningful play experience.

3.2 Current research

Deterding (2014) presents a need for rethinking gamification by stating that in only four years the whole concept has evolved from being almost a non-thing to a huge phenomenon that seemingly everyone wants a part of. However gamification is often reduced to a system for collecting points with no real connection to the tasks that are being gamified. Psychological, ethical and playful viewpoints are largely disregarded.

Gamification is a form of persuasive design, which means that it is also subject to the ethical discussion surrounding the topic. Deterding (2014) discusses the ethical issues from a practical viewpoint. Gamification as such is not inherently ethical or unethical, but it can be used in both ways. It should not be used, as it in most cases currently is, as a technical solutions for improving performance, but rather as a tool for creating interesting and motivating environments for work. On that note, Deterding (2014) says that “gamification needs to be rescued from the gamifiers.” Using lessons learned from game design it would be possible to build gamified systems that do motivate people and provide meaningful experiences that genuinely engage users.

Foursquare is probably the best known current platform for gamification. It has brought a certain form of gamification virtually everywhere. Foxman (2014) analyses it both from personal experience and existing research. It should be noted that Foursquare has subsequently changed their model of operations slightly, but for the most part Foxman's findings still apply.

Foursquare has been criticised for promoting consumerism and manipulating users to frequent certain establishments, just as Deci et al. (2001) said controlling rewards do. This may be true as such, but does not tell nearly the whole story about the service or users. It still managed to gather a considerable user base, though some of it may have been due to the novelty of the idea. Currently the service is undergoing an interesting change, with an attempt to encourage interaction between users with for example user tips. The changes at least on the surface seem to lessen the level of commercialism and promote activity in the community, and in some sense move from controlling rewards towards informational. In any case user behaviour seems to support what Schrape (2014) discussed about people being willing to give up private information. Foursquare users are very enthusiastic to give up their privacy in form of location data and places they frequent in order to gain virtual points and badges, and to score higher than their friends on leaderboards.

The check in functionality of Foursquare is also used as a tool for “proximal communication” (Foxman, 2014), where users check themselves in to a certain place, primarily perhaps to collect points, but with the secondary goal of letting their friends know they are there. It could be argued that in a pervasive game like Foursquare there is a lusive goal of points, but also a prelusive goal of contacting people and possibly socialising with them. It could also be argued that in this case gamification is in fact changing the way people behave. This again brings up the questions of user manipulation and touches also what Schrape (2014) discussed about gamification offering a positive feedback loop that is more easily obeyed than control by rules and restrictions.

The ability of gamification to motivate people does not come solely from the promise of rewards. According to several studies presented by Raczkowski (2014), the abstract scoring system and points as such are enough to get users competing over them. Scoring systems may be used for sorting players with high scores, or to give them additional skills, levelling up. Points create meaning out of nothing, hence turning any activity into a personal game. There are of course many other reasons for playing games, but that does not lessen the importance of points.

But simple score keeping systems are not games. They are not ludic by nature and only serve as tools for motivation and measurement. In some cases gamifying a system in fact artificially creates scarcity of resources and limits player behaviour. Users of a gamified system are not seen so much as players, as they are test subjects, who are only there to serve a purpose, not to engage in play. (Raczkowski, 2014)

Nicholson (2014) presents is an interesting comparison between MMORPGs and gamification. Specifically he discusses the element of grinding in both activities. Grinding is the act of performing repetitive tasks in an RPG, the goal of which is to collect experience points and eventually improve the character by gaining levels. In the scope of gamification the ordinary and

mundane task that is being gamified can be seen as grinding, something that has to be done in order to reach a goal. Nicholson presents a four step system that most MMORPGs follow:

1. Do quests and slay monsters.
2. Gain experience points.
3. Level up, become more powerful.
4. Move to new, more challenging areas and start over.

In comparison, the system for gamification works somewhat similarly. Doing tasks (1.) leads to gaining points and badges (2.) which leads to not only moving up on leaderboards, but also to gaining useful skills in real life (3.). What Nicholson says is lacking from most gamification implementations however is the fourth step. There are no new areas to move on to and users are stuck with grinding the same tasks and gaining the same badges seemingly indefinitely. This is particularly a problem with long term engagement, where rewards can replace the original intrinsic motivation with extrinsic one. If the gamification system is then discontinued, it may leave users unmotivated (Zichermann & Cunningham, 2011, p. 27). However, as can be seen from the analysis conducted by Hamari, Koivisto and Sarsa (2014), gamification does seem to work in bringing positive results in short term.

The comparison of MMORPG's and gamification goes on with the notion of endgame. That is the part of game when there is no more reason to grind. In other words the character has been developed as far as possible. At that point games often present the players with even more challenging tasks in order to keep them engaged. This seems often difficult or impossible in a gamified system. This would again seem to support what for example Deci et al. (2001) said about extrinsic motivation being harmful, since it can lead to a spiral of users wanting more rewards for the same tasks.

Nicholson goes further in a paper where he suggests several methods for creating meaningful gamification, as oppose to reward gamification (Nicholson, 2015). While rewards work well in the early stages of learning a skill, when the skill itself becomes meaningful, the rewards are no longer needed or even wanted. Rewarding can also be a functional incentive when there is little or no chance of creating intrinsic motivation for users. But even in these cases, rewards should be faded out as the process continues, otherwise users may grow accustomed to being rewarded and expect ever increasing rewards. In the solution Nicholson provides he suggest that game elements other than points and rewards are used in gamification.

The problem with this approach however is that the feeling of meaningfulness is highly subjective. As was mentioned earlier in reference to self determination theory (Deci et al., 2001), to which Nicholson also refers to, people find motivation in different ways. Therefore Nicholson's RECIPE-approach presents six aspects of meaningfulness.

1. Play, which by definition is an optional activity. Therefore any gamification system must allow free exploration and learning by doing.
2. Exposition, or the narrative layer, where past, present and future can be tied together in a meaningful way. It can also connect real world to the gamified environment with for example analogies.
3. Choice, users must have the ability to decide for themselves how to use the game elements, even if it means not using them at all. This relates closely to the voluntary nature of play. The downside here is imposing a lot of responsibility to the user, so allowing a free setting of goals and then offering guidelines on how to reach them could be a more feasible approach than total freedom.
4. Information, how and why points for example are given, and how rewards relate to the real world. Information can be conveyed with graphics or NPCs within the game world.

Deciding which information is important is however both subjective and situational, which makes it difficult to provide the correct amount of information at the right moment.

5. Engagement, both socially with other players and mechanically with the game world. This links to the relatedness aspect of self determination theory, and to the concept of flow. It should be noted that for many users reaching social engagement requires a fair amount of skill in the game. People are not willing to go and play with others if they feel they are inferior, so it would be important to allow practice. Social interaction can be reached either through competition or cooperation, or in some cases both. Users can be divided in competing teams for example. Cooperation can also happen in unexpected and unintended ways. Especially in a networked environment people are likely to ask for help in difficult places and share tips on possible loopholes in the system.
6. Reflection, which allows users to understand the real world implications of their game world actions, which deepens learning. Reflection works better in groups and often requires at least some guidance, but is a very powerful tool especially when looking for permanently changing behaviour, like in a fitness applications for example.

Overall, when it comes to Nicholson's view on meaningful gamification, it should be seen as a collection of layers that build intrinsic motivation towards an activity and can be peeled off one by one. In the end there should be no longer a need for motivational tools other than the activity itself, somewhat similarly to what Csikszentmihalyi (2013) suggested about making workers aware of the importance of their own work instead of giving them abstract rewards.

Currently, when designing gamification, games are generally seen as closed systems of their own. The complex interactions they provide are largely disregarded, which leads to games being seen as collections of mechanics. What Deterding (2014) suggests is an approach to gamification where

context, rather than game design elements, are central to gamification. In this way of thinking social norms are considered to be central in controlling behaviour.

First of all autonomy and voluntary participation are deemed important aspects of games. Games are not played voluntarily because they are fun, but rather the other way around. As already mentioned, autonomy is an important part of intrinsic motivation (Deci, 2001), and adding external rewards to that system reduces motivation.

Secondly situational norms dictate what is and is not acceptable. The social context of games allows for a certain amount of selfishness in order to win. There are of course rules limiting what can be done in order to maintain the aspect of fun in a game, but generally speaking increased amounts of selfish behaviour are tolerated in the context of games. Another crucial aspect of games is that they can be played against their rules in ways the designers did not intend. These aspects are usually lacking from gamification and the problem can be increased by rigid structures. Large organisations attempting to gamify their procedures often fail to see the need for a fundamental change and only manage to absorb gamification to be a part of their already existing systems.

Thirdly embarrassment is an important aspect of social behaviour. People often behave in certain ways in social situations to avoid embarrassment. Play is seen as an activity of leisure and pass time and is not socially acceptable behaviour in the workplace. Even in relation to so called serious games the social setting is not deemed playful, which reduces the enjoyability of those sort of games.

Philippette (2014) continues on the distinction between game and non-game activities. Certain activities are inherently seen as games and through that fun, while others are not. He also says that the form of gamification where points seem to be the main issue is inherent specifically to the world of electronic games. What Philippette proposes instead is that it is the spirit of a game that causes people to play and perceive an activity as a game, not the points and scores. He goes on to claim

that playability comes from various facets, as he calls them, which include aspects like game mechanics, interaction and artistic aspects. Furthermore, the experience of play is a subjective one. Things like previous experiences and social context affect the way that the facets of a game are perceived.

In the heart of every game however there is competition, either against a computer generated opponent or other human players. If this component exist together with a sufficient amount of perceived playability, an activity is accepted as a game and treated as such. In that sense Philippette seems to suggest that play comes naturally.

With gamification however, there are certain fallacies, as Phillipette puts it. Digitalisation fallacy suggests that many of the current gamification implementations are nothing more than glorified customer loyalty systems, which are not perceived as games. Roots of this idea are in the notion that gamification is done largely based on the mechanics of electronic games. Behaviouristic fallacy on the other hand suggests that cultural surroundings alone are not enough to condition people into perceiving activities as games. Merely calling something a game does not make it a game. Lastly, aesthetic fallacy says that visual characteristics alone do not transform an activity into a game.

Finally Philippette goes on to claim that the biggest problem of gamification is the game fallacy. According to that, it does not matter how game-like an activity is, if it does not get perceived as a game. Therefore in gamification designers should think like players, not like game designers.

Continuing in a way with what Philippette (2014) said about games and non-games, Dragona (2014) draws an interesting picture of gamification and the magic circle, as proposed by Salen and Zimmerman (2004). In her paper she asks if pervasive gamification is at all feasible, as it seems to counteract the magic circle by being everywhere. However there is also a suggestion that a lusory attitude, in turn proposed by Suits (1978), would be in the rise overall in society. This would in part explain the increased interest and willingness to accept gamification. This was also discussed by

Raessens (2014), saying that there indeed are some cultural indicators suggesting a change towards the playful. Specifically gamification is in part used to encourage data gathering in movements like the Quantified self,⁵ where data is gathered through a variety of sensors to create a comprehensive picture of individual achievement. Another interesting point is what could be called unseen gamification that goes on in for example social media, where likes and shares can be seen as a scoreboard of sorts with competition over virtual credit. Subsequently, some social media services are also played in ways that were not intended, meaning that playful behaviours find their ways through sets of rules, as they sometimes do in games as well. What seems to remain unclear however is how these activities are perceived by users and how they fit within the problematic characterisation proposed by Philippette (2014).

Fizek (2014) discusses the idea of emergent playfulness, by which she means the aspects of games that come not from intended gameplay, but from forms of play that are created by the players in unplanned and unexpected ways. Where Philippette (2014) talked about playability being subjective and somewhat difficult to come by, playfulness only emerges from the possibility of free movement. Notably, it is not important if the free movement take place within the game or with it, as an outside actor. This would suggest that activities that are not inherently playful can be seen as such, if they allow enough outside interaction through creating competition or win / lose conditions.

4 METHOD AND TARGET GROUP OF STUDY

The objective of this work is to collect information about how users of a gamified self-learning platform Duolingo perceive the effects of gamification. This chapter will describe Duolingo as a service, the selection of method, formulating the questionnaire and the limitations of this approach in general and in this particular case.

5 <http://quantifiedself.com/> (Visited 2.3.2015)

Based on what has been discussed above in chapters 2 and 3 the research questions were formulated to the following form:

1. Does Duolingo as a service support the psychological dimensions as described by Kahn?
2. Is there a significant undermining effect for intrinsic motivation?
3. Does Duolingo encourage long term learning?
4. Does using Duolingo induce a state of flow?

Two of the first ones can be considered to be the primary research questions, while the latter two are additional questions aimed towards providing supportive information.

4.1 Duolingo

Duolingo is an online self-learning service that is aimed at people who want to teach themselves languages. The service is free of charge and available both as a web service and an Android, iPhone or Windows Phone app. For the scope of this paper it is not important which platform users were using, since the reward system is roughly the same throughout all of them. In short, users go through lessons in their target languages and each lesson consists of about twenty questions at minimum. The number of questions varies, because incorrect answers do not take users closer to finishing a lesson. After each lesson users are rewarded with experience points (XP) and with enough points users gain levels. Successfully finishing lessons also opens the next level of lessons, which are otherwise unavailable. Users can also set a daily target of XPs they want to achieve, and achieving that target on 5, 10 or 20 consecutive days (a streak) is additionally rewarded. There is also a reward system called Lingots, which are a sort of virtual currency. They are rewarded for levelling up, completing streaks and finishing lessons without mistakes. Lingots can be used to purchase items from a store, there are for example special lessons such as flirting and idioms that can only be purchased with Lingots. Gambling is also possible, user can purchase a double or

nothing card for five Lingots which are doubled if the user finishes a five day streak. It should also be noted that the version studied here is the one aimed at regular users. There is also a version of Duolingo aimed for teachers, where it is possible to create customised courses for pupils to work at.

The reason for selecting Duolingo and its users as the target of the study was that it uses what could be called a traditional selection of game-like elements in order to engage users and it seemed fairly straightforward to formulate a questionnaire to measure user perceptions. The user base as such is also relatively wide and internationally distributed, which was assumed to provide a good response rate. In addition the service is relatively simple, which allowed breaking the structure down to core elements that could be analysed in light of the research presented above.

4.2 Selection of method

The questions that are being measured are subjective in nature and thus present some obstacles for data gathering. As for example Taylor-Powell (1998) says, measuring things like opinions, values and attitudes is difficult. The concepts are abstract and people usually cannot articulate them very well. Therefore the survey must break the concepts down to their core elements and ask about questions about those.

Online questionnaire was selected as a method for this study simply because the target respondents of the survey are not easily reachable by other methods. The reason for not selecting for example an email interview is that there was an initial idea about gathering data from a larger group of users instead of getting individual opinions. There are also some open questions in the questionnaire, which are used to deepen the understanding on some of the points. It is also important to note that the user base is not homogeneous, and it is distributed over geographical and cultural barriers. Therefore interviewing a relatively small number of people would not give a comprehensive picture of the entire host of users. There also was a desire to gather information specifically from the viewpoint of users, not that of the organisation. The primary focus is not to understand how

gamification can improve learning results, but rather how the users perceive it from a personal point of view.

A questionnaire study can be divided in three sections, which are data gathering, measurement and analysis. This chapter describes the data gathering. Measurement and analysis are presented in following chapters. The gathering stage includes both drawing up the questionnaire and distributing it among users, and it is considered to be the most challenging phase of the work. Mistakes made in the data gathering cannot be corrected later on, so a lot of attention needs to be put in the correct formulation of the questionnaire. Some major issues when drafting questionnaires are for example scope of the answers, correctness of the questions and reliability of selected measurements. All these issues needed to be considered when drafting the questionnaire. Vehkalahti (2008, p. 20) summarises the correct form of a questionnaire well by saying: “it has to ask substantially correct questions in a statistically meaningful way” (translated from Finnish).

While questionnaires are traditionally seen as tools for gathering quantitative data on a more general level (Vehkalahti, 2008, p. 13), they can also include open questions that provide qualitative data that goes into more detail. Open questions are also good for finding out attitudes and motivation behind actions, and can sometimes bring out some core points that would otherwise go unnoticed. Closed questions on the other hand are easier to compare and analyse, and also allow the respondent to use the questionnaire as an aid for recognition, rather than having to recall events without any support. (Hirsjärvi, Remes, & Sajavaara, 2009, p. 201).

Possible pitfalls of the method are things like low response rate and shallowness of data. As Crawford (1997, p. 33) states, careful thought should be put to developing a questionnaire in order to make sure the data gathered really answer the questions the study is interested in. Crawford also lists nine preliminary steps in questionnaire design that this work intends to follow.

1. Decide the information required.

2. Define the target respondents.
3. Choose the methods of reaching the target respondents.
4. Decide on question content.
5. Develop the question wording.
6. Put questions into a meaningful order and format.
7. Check the length of the questionnaire.
8. Pre-test the questionnaire.
9. Develop the final survey form.

In addition to the problems described above, this type of research poses other challenges in different stages of the process as well. In this particular case for example cultural differences can be an issue. As mentioned, the user base is distributed to almost every continent and cultural differences are likely to have at least some effect on for example preconceptions about games, the separation of play from work or work ethic in general.

While formulating questions in a correct manner is important, the questionnaire should also be distributed to a representative audience. In this case the recipients are all users of a certain learning tool, so representativeness is not an issue in that sense. Response rate however is often mentioned to be one of the biggest problems with surveys (Hirsjärvi et al., 2009, p. 195), so it was important to come up with an efficient way to distribute the questionnaire. To reach as many users as possible a link to the survey was posted on Duolingo's pages on Google Plus and Facebook, as well as through Twitter. After a week a reminder was also sent to boost response rate. The survey was open in total for two weeks.

The first draft of the questionnaire was formed fairly quickly and was largely based on research discussed in chapter 2. The questions were later revised to better suit the work at hand. This second version was then tested by three people who are not a part of the target group but were still capable of providing some useful insights regarding for example wording, understandability and the order of questions. Based on their suggestions a third version was formed and eventually distributed to the recipients. The final version of the questionnaire was implemented using the Form tool of Google Drive and is presented in appendix 1.

4.3 Formulating the questionnaire

First of all, it was important to determine which elements need to be present in order to create an engaging experience. Listed below are the concepts that emerged from the research described in chapters 2 and 3.

1. Tangible rewards, which are seen as controlling the user and providing extrinsic motivation.
2. Feedback, or verbal rewards, which can be used to encourage users to better results.
3. Companionship, in form of collaboration.
4. Competition against other users or the system itself.
5. Support in difficult situations.
6. Exploration, autonomy and problem solving.
7. Clear goals.
8. Creativity.

Table 1 depicts the questionnaire claims and table 2 the open ended questions. The process and reasoning behind the questions is described below.

1	Lesson goals are clear and achievable.
2	There is autonomy in the lessons.
3	Lessons are difficult enough.
4	Users can choose their own path through the program.
5	There are unexpected tasks.
6	Lessons can be finished easily.
7	Finding answers requires creativity.
8	There is enough support to get through difficult tasks.
9	There is positive feedback on correct answer.
10	There are sufficient instructions on incorrect answers.
11	The reward system (XPs and Lingots) is encouraging.
12	Lessons are easy to remember afterwards.
13	Rewards encourage returning to the program.
14	There are incentives to perform well.
15	The program supports trial and error.
16	Incorrect answers are handled in a supportive way.
17	The program removes distractions.
18	There is a sense of trust towards the system.
19	There is a chance to compete against other users.
20	There is a chance to collaborate with other users.
21	Users can utilise learned skills in practice.

Table 1: Questionnaire claims on a 1-5 Likert scale from strongly disagree to strongly agree.

1	In your opinion, what are the greatest benefits of Duolingo compared to using other learning methods?
2	In your opinion, what are the greatest drawbacks of Duolingo compared to using other learning methods?
3	Have you used the skills learned in Duolingo in a real life situation? If so, how and where?
4	If you have used other self-learning tools similar to Duolingo, how would you compare them?
5	If you have any other opinions or comments about Duolingo, please feel free to express them here

Table 2: Open ended questions.

Psychological dimensions (Kahn, 1990) that were discussed earlier have an important role in user engagement and special emphasis was placed on them when designing the questions. The dimension of psychological meaningfulness was fairly simple and questions about challenge, clear goals and autonomy also come up with other sources (e.g. Deci, 2001; Csikszentmihalyi, 2014). Psychological safety was also well represented in other sources and simple to form into questions about support and fear of losing. In terms of psychological availability, web-based systems are somewhat difficult since so many subjective factors come to play. Problems in personal life for example may affect the perceived availability, but asking about personal life in a questionnaire can easily be seen as overly intrusive. Similarly, since Duolingo is mainly used at home, the question of distractions is also difficult to interpret correctly.

As was mentioned above, most gamification systems seem to revolve around awarding the user with some abstract and seemingly unrelated rewards in order to maintain interest. These rewards were however discussed at length by (Deci et al., 2001), largely concluding that such rewards are not beneficial from the viewpoint of the user. Only if the task at hand is inherently boring, external rewards can have some positive effect. In addition Csikszentmihalyi (2014) states that instead of rewarding people for doing boring tasks, the importance of those tasks should be made plain to them.

As Deci et al. (2001) divide rewards to verbal and tangible, feedback can also be seen as a form of reward. In this case of course the word verbal must be extended to also include written feedback. Tangible rewards in turn are extended to include virtual rewards, such as points and badges. As was discussed above, informational verbal rewards can be seen as non-controlling, in other words they do not alter user behaviour, whereas tangible rewards are more likely to be controlling. Feedback however is an important part in reaching a state of flow. It would seem then, that providing sufficient feedback would be important for both encouraging the user and directing towards a flow state, while they do not undermine intrinsic motivation as tangible rewards seem to do.

In the questionnaire rewards were divided according to the division above, but terminology was changed so that verbal rewards are referred to as feedback and tangible rewards simply as rewards. Tangible rewards were studied according to the division of task non-contingent, task-contingent and performance-contingent. In other words users were asked if they feel like they are getting rewards for merely showing up, so to speak, for doing some tasks or performing well in tasks.

Companionship is an important factor within a social hierarchy, and it is also important in terms of engagement. When two equal peers for example work on a problem, they can share their experiences and learn from each other. It is more likely that people working in collaboration invest themselves in the activity, because they do not wish to let anyone down. Having a system that supports collaboration and companionship would be important in terms of increasing user motivation and engagement, so a question about companionship was also included in the questionnaire.

Competition is very interesting factor from the viewpoint of engagement and motivation. On one hand it would seem to contradict companionship as described above, but it is not necessarily always the case. The most obvious case of this is a team effort, where one team competes against another. In this setting there is strong companionship within a team and strong competition between teams.

The problem with competition does not end there however. Csikszentmihalyi (2014) points out that there are for example some zero-sum games, where in order to be a winner there is not necessarily a loser. Many single player electronic games for example fall under this category, as do solitaire card games. But even these sorts of games can be turned into competitive games by adding a score system, which of course resembles external rewards as described by Deci et al. (2001). Furthermore, as winning seems to be an important factor in competition, there must also be the option of losing. This in turn seems to hinder reaching a state of flow, because in order to reach it, one must be in a position where there is no fear of losing (Csikszentmihalyi, 2014). The majority

gamified systems seem to revolve around adding points and external rewards to existing tasks and in some cases promoting competition between users. This seems to be completely opposite to what would be required to truly engage users, instead they of just hook users to a scoring system for a quick reward.

Free exploration is often described as one of the core elements in games and play. In gamification however, this aspect often gets lost. Finding the unknown and being able to solve problems in novel ways are also very important aspects in learning and promote a positive self image. The investment of preferred self can also be increased when a person is presented with a novel task that is not previously known (Kahn, 1990). Flow theory seems to contradict in a certain way the idea of exploration by stating that in order to reach a state of flow, a set of clear goals must be set. This does not seem to go well with free exploration. However it could be argued that in a gamified system, the user could be presented with a clear overall goal and then be left free to find a path to it, perhaps only providing some guidance on the way if necessary.

4.4 Limitations

The reliability of a an online questionnaire is always a question. In this case for example there is no way of knowing for example what the percentage of responses is. Only estimates can be made based on the amount of users Duolingo report they have, which does not take into account the level of activity. Furthermore there is no way of knowing what percentage of the user base received invitations to fill the survey, thus there is no data on how actively the survey was filled. It is also worth noting that while the survey is targeted to a very specific group of people, it is not known for sure if a certain subgroup is over presented.

Vehkalahti (2008, p. 40) states that surveys are often the result of cooperation between experts in different fields. At the very least he says that a statistician should be involved already in the drafting phase, together with an expert in the field in question. This being a master's thesis work that is not

the case and may result in compromising the validity of the questionnaire. In other words even after careful formatting and testing there may still be issues with the questionnaire asking the correct questions. Another validity issue that Vehkalahti mentions is that cultural and language may play a part when conducting surveys with questionnaires. In this case the questionnaire was implemented in English, which is not the native language of the surveyor, nor is it the native language of all the respondents. Even though Duolingo itself relies heavily on users speaking English, there is still a chance of misunderstanding.

The greatest limitation in this case however seems to be the low response rate. Even after promoting the survey on social media pages targeted towards Duolingo users, there were less than ten responses. Subsequently, a message was posted to a private Facebook page of the author, which managed to boost response rate up slightly, but still at the end of the two week period there were only 26 responses. At this stage the decision was made to close the survey in any case, since it seemed unlikely there would be any more responses and there was a time limit on how long the survey could be kept open.

5 RESULTS

As was mentioned in the previous chapter, the response rate was lower than expected. In the end of two weeks being open, there were only 26 responses, 11 female and 15 male. Due to the low response rate the results are not conclusive, but they are presented below nonetheless. Ages varied between 15 and 74, with the mean age of 32 and median of 33. Geographically the respondents were strongly focused in central and northern Europe, with only a few additional responses from North-America. Native languages are displayed in table 3. The skill level of recipients was determined by asking what their highest level in Duolingo is. The mean for user skill level was 9,9 and median 8, with a standard deviation of 7,5.

English	12
Finnish	7
German	2
Danish	1
Dutch	1
French	1
Spanish	1
Swedish	1

Table 3: Language distribution

As mentioned, the survey was initially distributed via social media to Duolingo users, but due to a low response rate it was also promoted through the authors personal Facebook page. This is probably the reason for the centralisation in geographic and age. In spite of the low response rate, there was some saturation in the open ended questions, which were answered relatively well.

To answer the presented research questions the questionnaire data will be divided to respective subcategories and is discussed accordingly.

5.1 Psychological dimensions

To test psychological meaningfulness, there were questions concerning goal clarity, creativity and task autonomy. Combined 76,9 percent of the respondents said that they either agree or strongly agree with the claim that lesson goals are clear and achievable. In task autonomy and the ability to choose one's own path however, the results were almost opposite with 65,4 percent of the respondents disagreeing or strongly disagreeing on autonomy and 69,3 percent disagreeing or strongly disagreeing on the ability to choose a path. There is a similar story with the question concerning creativity, where 77 percent of respondents disagree with the claim that finding answers requires creativity and only one respondents agreeing with the claim (Illustration 2). Overall it seems that even though users understand what their goals are, there are only few predetermined ways to achieve them and free exploration is not supported.

In the open ended questions however the ability to set own personal goals and work at one's own pace to achieve them was mentioned on two occasions. There were also several mentions on how the systems is not based on strict rules and that it imposes no rigid structure to follow. However the ease of use and convenience were mentioned on ten occasions, which would seem to indicate that rather than being engaging, the system is merely convenient, offering simple tools for learning when and where the users desire.

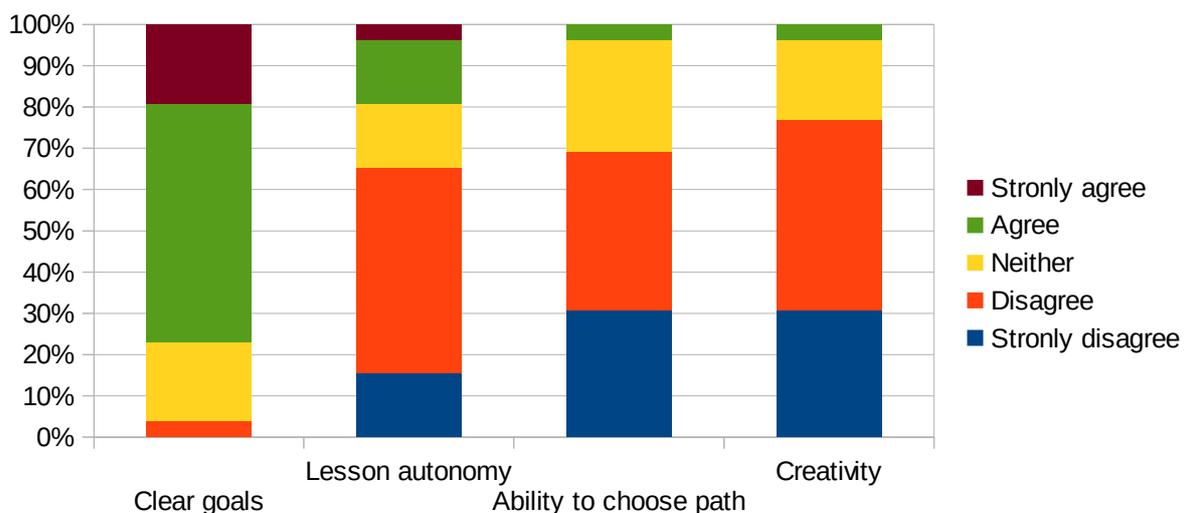


Illustration 2: Goals, autonomy and creativity.

In the question of psychological safety, aspects like supportive and non-threatening nature of the environment were measured. In the question of overall support in the tasks, 57,7 percent of the respondents agreed or strongly agreed with the claim. In the question of support with incorrect answers however there was a near equal distribution between disagree, neither, and agree. Learning through trial and error, which also requires users to trust that the system will support them in the case of the latter, 57,6 percent of respondents disagreed or strongly disagreed (Illustration 3). This would seem to indicate the even though there is some general sense of support, in individual tasks users cannot rely on getting instructions and not being left alone in trying to cope with mistakes. The open ended questions would seem to support this idea, since there were several mentions on the lack of support, particularly in relation to grammar. There seemed to be a desire for better instructions to begin with, and especially for feedback on what went wrong. When asked about the

greatest drawbacks of Duolingo one recipient for instance wrote: “Sometimes it's difficult to understand why a response was wrong.” In relation to psychological safety, this does not seem to support the claim that Duolingo would be engaging in this manner. Even though lessons are easy to go through, learning seems to happen more through repetition than understanding.

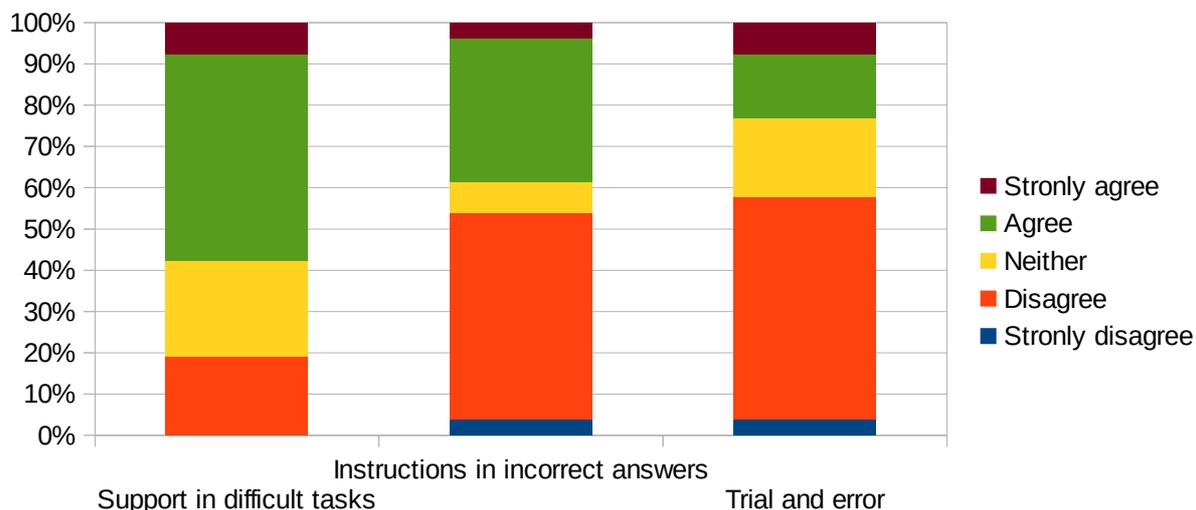


Illustration 3: Support and instructions.

For psychological availability the questions were difficult, since it is perhaps the most subjective of all the dimensions presented by Kahn (1990). Questions concerning personal life were not asked in the questionnaire, and finding out how users perceive the systems ability to remove distractions probably depends more on the environment in which each individual uses it. When asked about the programs ability to remove distractions, 65,4 percent of respondents disagreed or strongly disagreed with the claim, but it should be noted that at the same time 19,2 percent could not state their opinion. Combined 80,8 percent however agreed or strongly agreed with the claim that there is a sense of trust towards the system (Illustration 4). What is particularly interesting with these answers is that again in the open questions two respondents, both native English speakers from North-America, claimed that there are some poorly formulated English examples, and even outright errors in grammar. One pointed out that there is no way for a language learner to know if a translation is correct. Even in spite of that, most recipients seemed to be happy with the systems and trusted it.

This would seem to indicate that at least based on the available data, perceived availability supports psychological engagement.

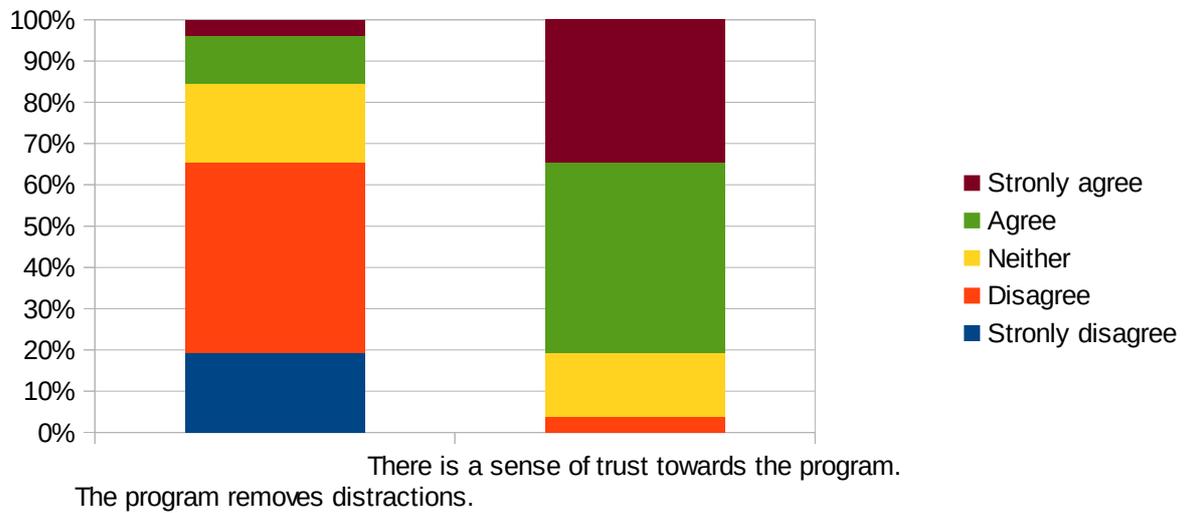


Illustration 4: Distractions and trust.

Overall what has been discussed above would seem to indicate that as far as user engagement is concerned in the sense that Kahn (1990) described it, Duolingo as a learning platform does not utilise tools that would strongly engage their users. Ease of use and convenience seem to be more important as reasons for using the system, and even though there seems to be some evidence about overall support and trust, there is no guidance for users in lessons. This leaves uncertainty about how to correct answers and what to do next.

5.2 Rewards and motivation

The questions about rewards were divided to verbal rewards, referred to as feedback, and tangible rewards. The informational – controlling axis used by Deci et al. (2001) is also used to find whether or not the reward system used by Duolingo is likely to alter user behaviour and replace intrinsic motivation.

When asked about feedback on correct answers 42,3 percent of the respondents could not either agree or disagree and this aspect must be deemed inconclusive. Support on incorrect answers was reported above to spread evenly across agree and disagree, and support on correct answers spreads

similarly. This would seem to indicate that informational verbal rewards are not very well supported by Duolingo.

For tangible rewards, 88,5 percent of respondents considered the reward system encouraging with 84,7 saying it makes them to come back to the program. Only 46,1 percent however said they feel there are incentives to perform well (Illustration 5).

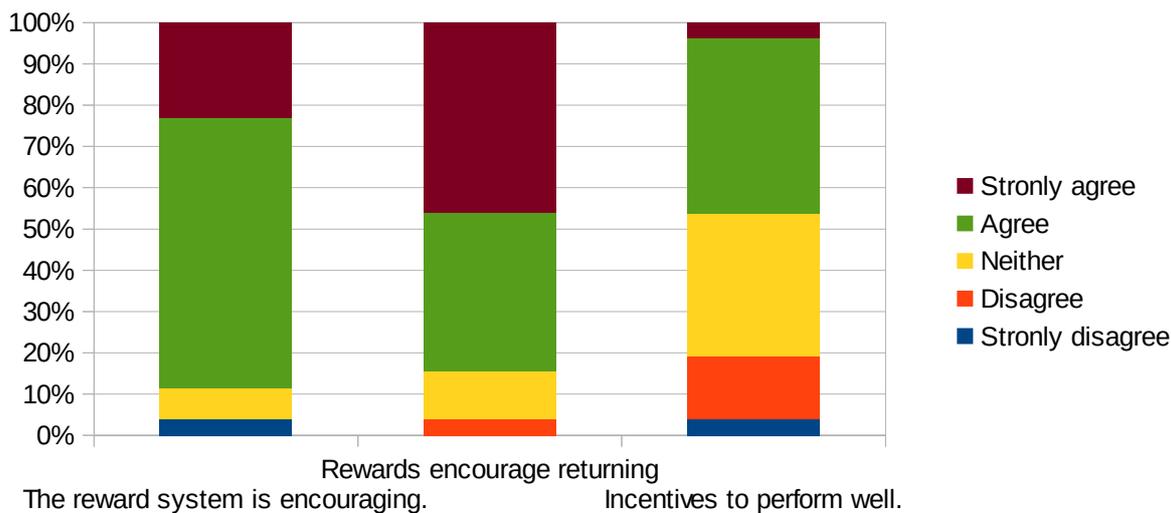


Illustration 5: Tangible rewards.

The question of long term learning is also one that has a connection with both verbal and tangible rewards, since increasing extrinsic motivation is considered to undermine the effectiveness of lessons. When asked if the users find it easy to remember the lessons afterwards, the results show that while 42,3 percent either agree or strongly agree, 30,8 percent cannot say and 26,9 percent disagree or strongly disagree. What seems to be particularly interesting in the question about long term learning is that there is slight negative correlation (-0,0077) between the reward system being encouraging and users remembering the lessons afterwards. On the other hand there is a more significant positive correlation (0,1971) between incentive to perform well and remembering the lessons. The open ended questions however there was some evidence to suggest fairly good results on long term learning as well. When asked if users have used the skills learned in Duolingo in a real life situation, 11 of the 26 respondents said that they had used a learned language at least in some

situation. One respondent claimed to be still in too early stages, but would probably use the language after progressing further. Three respondents denied having used their learned skills and the remaining 12 did not answer, which is probably interpretable as a no. Most respondents who had used what they had learned had engaged in light conversation with native speakers of the language, either on a holiday or in a work environment. Three had also read newspapers or books in their target languages.

The open ended questions would seem to support the conclusion that there are more incentives to participate than to perform well. When asked about the greatest benefits of Duolingo compared to other self-learning methods, one respondent simply said: "It's addictive." The rewards being a strong incentive were also mentioned in several other responses, both in terms of motivating users to come back and in the sense of creating a feeling of achievement. This seems like a fairly strong indication towards engagement-contingent reward system, which is the most controlling form of tangible rewards. In other words users are prone to alter their behaviour in the hope of receiving rewards, rather than for learning.

5.3 Flow

If a state of flow is seen as an extended state of engagement, it is also interesting to see how it manifests itself in the responses. Some of the questions concerning flow are the same as those with engagement, but users were also asked directly about the perception of their skill level and the ease of completing tasks. Overall 65,4 percent agreed or strongly agreed with the claim that lessons can be finished easily, while the same percentage also said that the lessons are difficult enough (Illustration 6).

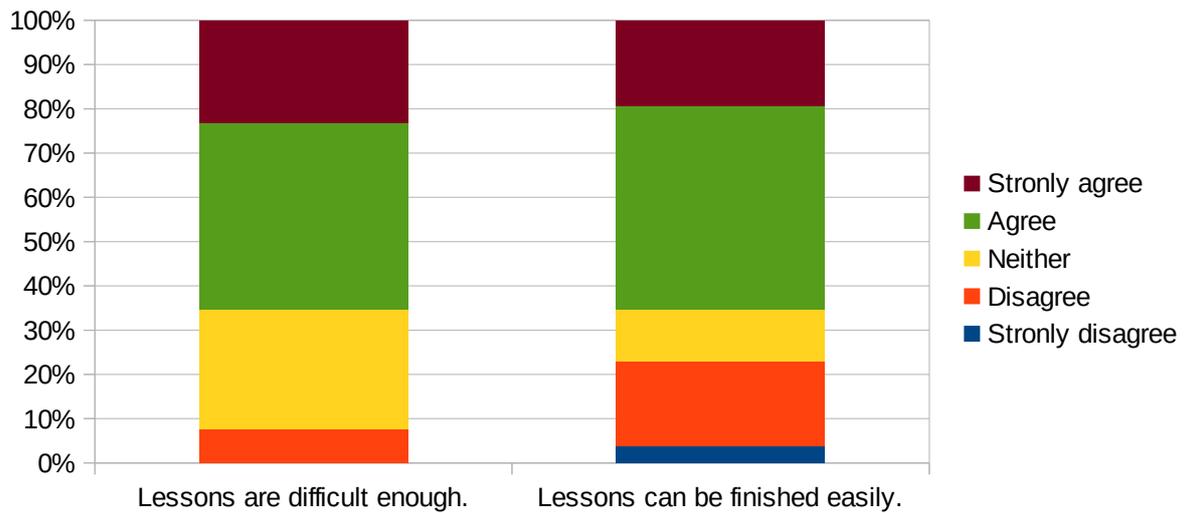
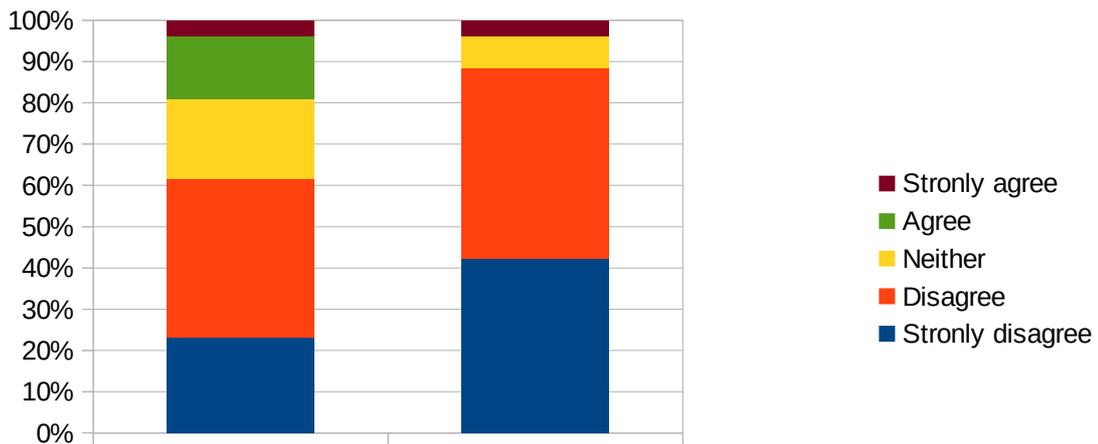


Illustration 6: Lesson difficulty.

Competition and collaboration, which are also seen as an important part in inducing flow experience, were however seen as not being present, 76,9 percent disagreeing with the claim that there is a chance to compete and 88,5 disagreeing with the chance to collaborate (Illustration 7). In terms of flow then it seems there is a division between users being satisfied with the way Duolingo lessons are crafted, but cannot use their skills to compete or to collaborate, which would increase the chances to reach a state of flow. In the open ended questions there were some mentions towards competition, one respondent said that the score system enables competition even without opponents: “I like the chance to compete against myself [...] Better than books at keeping up the interest.”



There is a chance to collaborate with other users.
 There is a chance to compete against other users.

Illustration 7: Competition and collaboration.

There also seems to be a correlation between user skill level and the perception of lesson difficulty. There is a negative correlation of -0,48 between the question “Lessons are difficult enough” and user skill level and a positive correlation of 0,44 between “Lessons can be finished easily” and user skill level. This is not surprising since it can be expected that more experienced users find lessons easier than novices. These data however do not provide an answer to the question of whether or not users perceive their tasks as routine, too difficult or indeed just at the correct level compared to their skills. There would seem to be some indications towards that, but it cannot be concluded. In the open questions there was critique particularly towards the grammar and in some cases mentions of errors. It would seem the user skill level has no effect on the criticism. Also those who feel there is a lack of support and examples on grammatical issues do so regardless of skill level.

6 DISCUSSION

Comparing the gathered data to the findings of current research described in chapter 3 seems to indicate that Duolingo does exactly what the research is warning about. It does not seem to induce user engagement, there is fairly strong evidence of that it offers controlling rewards, thus altering user behaviour, and does not seem to support reaching a state of flow. It would seem that the prelusory goal of completing a task is present in Duolingo. What may or may not be present for

each user is the illusory goal of playing a game. Similarly, it looks as if free play, labelled as *paidia* by Caillois (1958), is missing from Duolingo which is governed by a more strict set of rules.

It would seem that Duolingo offers precisely the kind of positive feedback loop that was discussed by Schrape (2014) and deemed as a good tool for altering user behaviour, but not for offering meaningful content. Rewards are described as addictive and strong incentives to come back, while useful feedback does not seem sufficient to support learning. Duolingo seems to have fallen to the trap of game fallacy (Philippette, 2014), offering fairly little in terms of elements that would normally be seen as game-like or playful. Instead there are points, levels and badges that are collected indefinitely. Even though each level seemingly presents users with new challenges, the nature of them remains much the same. This sort of grinding may quickly cause a gamified system to lose whatever novelty effect it initially has and cause users to become bored, or leave them wanting increasing amounts of rewards (Nicholson, 2014). This, according to Deci (2001), can have a negative effect on intrinsic motivation and through that long term learning.

In most gamified systems there is a fairly obvious motive for adding the layer of gamification. Like for example in customer loyalty programs, where the organisation is the beneficiary maintaining good customer relations with a relatively small investment. In the case of Duolingo however, the question is not so simple. As far as can be told, Duolingo does not charge private customers, nor does it advertise on site. So there are no obvious financial benefits that the company would gain from hooking up consumers to their service. That being said, it would still seem that the lack of meaningful content from a user perspective goes well together with the idea of a service only existing to make users into subjects, instead of actual players of a game (Raczkowski, 2014). So even though there is no obvious gain for the organisation, it still does not mean that gamification would be beneficial to the user.

Even though the gathered data does not as such reveal how long individual users have been using Duolingo, judging by the user skill levels the novelty effect must have worn off at least for some of the respondents. This would seem to indicate that even in spite of becoming routine-like, Duolingo still manages to maintain its appeal. In addition, gamification as such has been found to work well especially in the early stages of learning. There is no data available on how far Duolingo users take their language skills using the software, but there is some evidence in the gathered data that it has in some cases been enough to read literature and converse fluently with native speakers. There is of course no evidence about what other methods of education each individual uses, but it is reasonable to assume that as a starting point Duolingo works well. The overall increase of lususory attitude in society (Dragona, 2014) may also be one explaining factor to the popularity of Duolingo and other gamified services. Users may be more accepting towards game-like elements even in contexts where they are not traditionally seen, and do not necessarily even see them as game-like.

The social context of games is something that is quite different from the social context of work or other serious activities. Duolingo as a service is different from gamification systems that would be used for example in a workplace, because it is designed to be used at users own time, primarily for self improvement or even leisure. But even then, the point of it is to educate, and education is traditionally seen to be different from play. If the system is perceived as overly playful, it may lose some of its educational power. It seems like a difficult balancing act to incorporate enough playfulness to utilise the engaging aspects of gamification but still maintain enough seriousness to invoke serious learning.

As was mentioned by for example Deterding (2014), most people tend to avoid embarrassment in social interaction. In the case of language learning this is not necessarily a big problem, since languages are primarily learned through practice. There are bound to be mistakes and some increased level of embarrassment is accepted by all. Games are also a field where mistakes are not seen as a serious problem, since it is only a game. In the case of Duolingo there is limited support

towards learning through trial and error, possibly due to the lack of positive feedback, so it would seem that this aspect of benefits that gamification could offer are not utilised.

At least some of the findings discussed in chapter 5 seem to support what Nicholson describes in his RECIPE-approach (2015). The point of this approach is to use gamification in a manner which gradually removes the need for gamification, bringing the skill to front. In a way Duolingo seems to support playful learning, but does so in a restricted manner, leaning towards ludus on Caillois' axis. Free play is almost completely non-existent. The voluntary nature and free exploration that are inherent in playful activities do not come out in gamification when it is designed in a way that for example Duolingo is. There are instead strict rules that do resemble a game, but many elements that are otherwise required from a game are missing.

The informational content of Duolingo was mentioned by several respondents, specifically in a sense that there is not enough information available. There seems to be a missing connection between activities and rewards, which can lead users to only pursue the rewards instead of aiming towards gaining useful skills. What also seems to be missing is the chance for users to reflect upon what they have been learning. Reflection is a powerful tool to deepen learning, but it is also difficult to implement in an automated online system.

In order to better understand what Duolingo has to offer in terms of gamification and engagement and how to make it better, some of what could be called elements of user engagement are collected below in table 4. The explanations have already been presented in chapters 4 and 5 so they are not further addressed here. What the usage of these elements in Duolingo seems to show is that the game-like elements are relatively shallow and could be, according to recent studies, implemented better.

Element	Description
Rewards	Experience points, gaining levels and virtual currency.
Feedback	Limited to telling whether an answer was right or wrong.
Companionship	Not supported, group tasks for example are not possible.
Competition	Friends can compare performance with each other.
Challenges	User can take on extra challenges, rewarded with virtual currency.
Exploration	Not supported, only linear progress is possible.
Goals	Well defined and strict.

Table 4: Elements of user engagement in Duolingo.

Table 5 in turn shows the same elements, but instead of describing how they are used, it shows how they could be used, when taking into account what for example Deterding (2014) and Nicholson (2015) have said about gamification. It is a short list of how to improve Duolingo functionality, but can also serve as a checklist for designing gamification. The points are explained in more detail below.

Element	Description
Rewards	More closely tied to the tasks and lessons.
Feedback	More feedback in general, especially with instructions on how to proceed.
Companionship	Chance to collaborate with others.
Competition	More obvious chances to compete.
Challenges	Challenging lessons, perhaps with the possibility for custom challenges.
Exploration	Less of a rigid path through to lessons, more free play.
Goals	More self-determined goals.

Table 5: Suggested usage for elements of user engagement for Duolingo.

The reward system resembles the task – points – levels -model that Nicholson (2015) presented. The rewards are not clearly attached to the tasks and can leave users uncertain about the consequences of their actions. Instead of a regular point system there could be a reward system that is more tied to the tasks and lessons. Like for example a dictionary or a phrase book that would grow after each successful lesson and incorporate the words and phrases taught in it. The size of the dictionary

would be easily comparable with other users, but it would not be just an abstract layer of points. It would also help with revision, which is not currently very well supported.

Feedback was mentioned in several of the questionnaire responses as something the system could do better. Positive feedback, or verbal rewards, were present on some level, but instructions on how to proceed in problematic situations and exactly what went wrong with user responses were not perceived as clear enough. Using more verbal rewards and tying them more closely to each task would increase the sense of trust users have towards the system. More informational feedback on the other hand would increase the motivational characteristics and possibly enable deeper understanding, instead of just learning by repetition. To do this would be fairly straightforward and could be achieved by merely adding more instructions to each question.

Companionship can be an important part of learning, especially in a situation where peers can teach each other. But even without the aspect of teaching the sense of collaboration can improve learning results and increase motivation. Lessons could be designed in such a way that they could be completed in collaboration online, at it's simplest by dividing a set of tasks between two or more users and requiring all of them to be finished in order to continue. If Duolingo were to introduce more complex game elements to it, collaboration could be supported even further with for example puzzle-like games where hints are given to several users who need to work together in order to finish a lesson.

Collaboration could also be implemented in a form of reflection, where users could discuss what they have been learning and possibly gain some points from that. There already are some discussion forums where users do talk with each other, but such discussion is not outright supported or even suggested.

Competition, even though it can be seen as the opposite of collaboration, is also an important part in motivation and reaching a state of flow. Competition can be invoked either between groups or

individuals, or even between user and the system. Duolingo does already support competing against the clock, but that is an extra feature which needs to be purchased with in game currency. There is also the obvious chance for users to compare points and levels with each other, but it is not strongly encouraged by the system. On the other hand this could be seen as a good thing, since there are also those who say competition is bad for motivation, but allowing a chance for a group of friends to form a league where they could see for example who gets the most game currency over a set period of days might have beneficial effects. Though perhaps implementing some of the other suggested design modification might be required for competition to work.

Challenges are somewhat supported since users can set goals for themselves, like for example requiring a certain amount of points in a day to maintain a streak. More custom challenges like that however might increase motivation by keeping user alert. Especially if there would be unannounced challenges. Pop quizzes have been used in traditional education and they seem to work. Similar surprise tests could be implemented in Duolingo as well. Users can currently challenge themselves with tests which are purchased with game currency and there are different levels of tests, but they do not take the user skill into account on a personal level. The correct level of task difficulty is perhaps the most important aspect of reaching a state of flow, and if the system could dynamically adjust the level of difficulty, this would be better achieved.

Free exploration is important or even essential to play and games. In gamification however, it is challenging to produce an environment that can be freely explored but still manages to fulfil whatever educational or other tasks it is designed to do. At a very simple level free exploration could be supported by for example allowing users to freely pick their tasks from a set number of choices instead of forcing a rigid path that everyone needs to follow. This could at least create an illusion of exploration, even if all tasks would have to be finished before moving on to a new level. And as has been discussed above, the perceived experience is in many cases the most important aspect.

Goals are important for both motivation and for a state of flow. Clear and achievable goals are of course present in any system that revolves around education, but on a user level it is not necessarily always obvious what the objective of each task is. In other words the perceived goals are not clear which may lead to users simply doing tasks without engaging in the underlying objective, in this case learning. Even though the questionnaire responses would indicate that goals are clear, a question remains about how well users in fact internalise them.

7 CONCLUSION

In the recent years gamification has been heralded as something that can inject meaning to tasks that are otherwise seen as unimportant or uninteresting, as a tool to engage employees to their work and as a cheap method of marketing, hooking customers to a service with small rewards. This thesis set out to find out how the users of one specific gamified service, Duolingo, perceive the effects of gamification on their engagement and motivation.

In conclusion and to answer the research questions presented in this paper it has to be said that for the primary question, does Duolingo induce user engagement, the answer is no. There are not enough engaging elements to truly have users engaging in a meaningful way. The secondary question, is there an undermining effect for intrinsic motivation, it would seem that there is. Given rewards seem controlling rather than informational, and even though some users reported having learned useful skills and used them in real life, they also said the reward system is what makes them come back. The additional questions of long term learning and a state of flow were somewhat inconclusive and it must be admitted that observing such abstract phenomenon is rather difficult through methods that were available. The results would seem to indicate that Duolingo does not support reaching flow, but that varies greatly from one user to another.

With all the research material that is currently available on gamification, it would be possible to build a gamified system that would be both meaningful and usable, and in addition beneficial to

users and organisation. In the case of Duolingo however, this does not seem to be the case. Even though there are certain individual aspects in the system that on the face of it would seem to support meaningful gamification, the overall conclusion must be that in this case the greatest benefits of gamification remain unachieved and the system relies on points and rewards without linking them to the tasks.

Studying user perspective is an important aspect of gamification research especially now when it is rising to become a more mainstream phenomenon. Understanding how to design gamification in a manner that makes it engaging without careening to a spiral of ever increasing rewards would be important not only to the organisations implementing it, but also to the users. Through user centered meaningful gamification design users could have the chance to learn new skills in an engaging manner or see the importance of their work however mundane it otherwise may seem.

REFERENCES

- Bogost, I. (2011). *Persuasive Games: Exploitationware*. Retrieved from http://www.gamasutra.com/view/feature/6366/persuasive_games_exploitationware.php
- Caillois, R. (1958). *Man, play and games*. (M. Barash, Trans.). Urbana (Ill.): University of Illinois Press.
- Crawford, I. M. (1997). *Marketing Research and Information Systems*. Food & Agriculture Org.
- Csikszentmihalyi, M. (2013). *Flow: The Psychology of Happiness*. Random House.
- Csikszentmihalyi, M. (2014). Play and Intrinsic Rewards. In *Flow and the Foundations of Positive Psychology* (pp. 135–153). Springer Netherlands. Retrieved from http://link.springer.com/chapter/10.1007/978-94-017-9088-8_10
- Deci, E. L., Koestner, R., & Ryan, R. M. (2001). Extrinsic Rewards and Intrinsic Motivation in Education: Reconsidered Once Again. *Review of Educational Research*, 71(1), 1–27. <http://doi.org/10.3102/00346543071001001>
- Deterding, S. (2014). Eudaimonic Design, or: Six Invitations to Rethink Gamification. In M. Fuchs, S. Fizek, P. Ruffino, & N. Schrape (Eds.), *Rethinking Gamification* (pp. 305–331). Lüneburg: Meson Press. Retrieved from <http://opus.uni-lueneburg.de/opus/volltexte/2014/14305/>
- Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: defining “gamification.” In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments* (pp. 9–15). New York, NY, USA: ACM. <http://doi.org/10.1145/2181037.2181040>
- Dragona, D. (2014). Counter-Gamification: Emerging Tactics and Practices Against the Rule of Numbers. In M. Fuchs, S. Fizek, P. Ruffino, & N. Schrape (Eds.), *Rethinking Gamification* (pp. 227–250). Lüneburg: Meson Press. Retrieved from <http://opus.uni-lueneburg.de/opus/volltexte/2014/14305/>
- Fizek, S. (2014). Why Fun Matters: In Search of Emergent Playful Experiences. In M. Fuchs, S. Fizek, P. Ruffino, & N. Schrape (Eds.), *Rethinking Gamification* (pp. 273–287). Lüneburg: Meson Press. Retrieved from <http://opus.uni-lueneburg.de/opus/volltexte/2014/14305/>
- Foxman, M. (2014). How to Win Foursquare: Body and Space in a Gamified World. In M. Fuchs, S. Fizek, P. Ruffino, & N. Schrape (Eds.), *Rethinking Gamification* (pp. 71–90). Lüneburg: Meson Press. Retrieved from <http://opus.uni-lueneburg.de/opus/volltexte/2014/14305/>
- Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does Gamification Work? – A Literature Review of Empirical Studies on Gamification. In *2014 47th Hawaii International Conference on System Sciences (HICSS)* (pp. 3025–3034). <http://doi.org/10.1109/HICSS.2014.377>
- Hirsjärvi, S., Remes, P., & Sajavaara, P. (2009). *Tutki ja kirjoita* (15th ed.). Helsinki: Kustannusosakeyhtiö Tammi.
- Huizinga, J. (1944). *Leikkivä ihminen*. (S. Salomaa, Trans.). Porvoo: Werner Söderström Oy.
- Kahn, W. A. (1990). Psychological Conditions Of Personal Engagement And Disengagement At Work. *Academy of Management Journal*, 33(4), 692–724. <http://doi.org/10.2307/256287>
- Köhne, E., Ewigleben, C., & Jackson, R. (2000). *Gladiators and Caesars: The Power of Spectacle in Ancient Rome*. University of California Press.
- Malone, T. W. (1982). Heuristics for designing enjoyable user interfaces: Lessons from computer games. In *Proceedings of the 1982 conference on Human factors in computing systems* (pp. 63–68). ACM. Retrieved from <http://dl.acm.org/citation.cfm?id=801756>
- McGonigal, J. (2011). *Reality is Broken: Why Games Make Us Better and How They Can Change the World*. Penguin Press.
- Nelson, M. J. (2012). Soviet and American Precursors to the Gamification of Work. *SSRN eLibrary*. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2115483

- Nicholson, S. (2014). Exploring the Endgame of Gamification. In M. Fuchs, S. Fizek, P. Ruffino, & N. Schrape (Eds.), *Rethinking Gamification* (pp. 289–303). Lüneburg: Meson Press. Retrieved from <http://opus.uni-lueneburg.de/opus/volltexte/2014/14305/>
- Nicholson, S. (2015). A recipe for meaningful gamification. In *Gamification in Education and Business* (pp. 1–20). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-3-319-10208-5_1
- Norman, D. A. (2002). *The Design of Everyday Things*. Basic Books.
- Philippette, T. (2014). Gamification: Rethinking “Playing the Game” with Jacques Henriot. In M. Fuchs, S. Fizek, P. Ruffino, & N. Schrape (Eds.), *Rethinking Gamification* (pp. 165–200). Lüneburg: Meson Press. Retrieved from <http://opus.uni-lueneburg.de/opus/volltexte/2014/14305/>
- Raczkowski, F. (2014). Making Points the Point: Towards a History of Ideas of Gamification. In M. Fuchs, S. Fizek, P. Ruffino, & N. Schrape (Eds.), *Rethinking Gamification* (pp. 140–160). Lüneburg: Meson Press. Retrieved from <http://opus.uni-lueneburg.de/opus/volltexte/2014/14305/>
- Raessens, J. (2014). The Ludification of Culture. In M. Fuchs, S. Fizek, P. Ruffino, & N. Schrape (Eds.), *Rethinking Gamification* (pp. 91–114). Lüneburg: Meson Press. Retrieved from <http://opus.uni-lueneburg.de/opus/volltexte/2014/14305/>
- Salen, K., & Zimmerman, E. (2004). *Rules of Play: Game Design Fundamentals*. MIT Press.
- Schrape, N. (2014). Gamification and Governmentality. In M. Fuchs, S. Fizek, P. Ruffino, & N. Schrape (Eds.), *Rethinking Gamification* (pp. 21–45). Lüneburg: Meson Press. Retrieved from <http://opus.uni-lueneburg.de/opus/volltexte/2014/14305/>
- Suits, B. (1978). *The Grasshopper: Games, Life and Utopia*. Broadview Press.
- Taylor-Powell, E. (1998). Questionnaire Design: Asking questions with a purpose. *University of Wisconsin Cooperative Extension*.
- Vehkalahti, K. (2008). *Kyselytutkimuksen mittarit ja menetelmät*. Helsinki: Kustannusosakeyhtiö Tammi.
- Zichermann, G., & Cunningham, C. (2011). *Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps*. O’Reilly Media, Inc.

APPENDICES

APPENDIX 1: The Questionnaire

4/15/2015

Duolingo Survey

Duolingo Survey

The purpose of this survey is to chart the engaging qualities of Duolingo. It will take approximately 15 minutes to answer.

The data will be used in a master's thesis work for the university of Tampere, Finland. All responses will be processed anonymously and no data will be handed to third parties. The survey is in no way affiliated with Duolingo.

Any questions or comments can be emailed to tamminen.juhani.a@student.uta.fi

*Required

Using the program.

Please answer the claims with 1 being strongly disagree and 5 being strongly agree.

1. **Lesson goals are clear and achievable.** *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	<input type="radio"/>	Strongly agree				

2. **There is autonomy in the lessons.** *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	<input type="radio"/>	Strongly agree				

3. **Lessons are difficult enough.** *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	<input type="radio"/>	Strongly agree				

4. **Users can choose their own path through the program.** *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	<input type="radio"/>	Strongly agree				

4/15/2015

Duolingo Survey

5. **There are unexpected tasks.** *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	<input type="radio"/>	Strongly agree				

6. **Lessons can be finished easily.** *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	<input type="radio"/>	Strongly agree				

7. **Finding answers requires creativity.** *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	<input type="radio"/>	Strongly agree				

8. **There is enough support to get through difficult tasks.** *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	<input type="radio"/>	Strongly agree				

9. **There is positive feedback in correct answers.** *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	<input type="radio"/>	Strongly agree				

10. **There are sufficient instructions on incorrect answers.** *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	<input type="radio"/>	Strongly agree				

11. **The reward system (XPs and Lingots) is encouraging.** *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	<input type="radio"/>	Strongly agree				

12. Rewards encourage returning to the program. *

Mark only one oval.

1 2 3 4 5

Strongly disagree Strongly agree

13. There are incentives to perform well. *

Mark only one oval.

1 2 3 4 5

Strongly disagree Strongly agree

14. Lessons are easy to remember afterwards. *

Mark only one oval.

1 2 3 4 5

Strongly disagree Strongly agree

15. The program supports trial and error. *

Mark only one oval.

1 2 3 4 5

Strongly disagree Strongly agree

16. Incorrect answers are handled in a supportive way. *

Mark only one oval.

1 2 3 4 5

Strongly disagree Strongly agree

17. The program removes distractions. *

Mark only one oval.

1 2 3 4 5

Strongly disagree Strongly agree

18. There is a sense of trust towards the system. *

Mark only one oval.

1 2 3 4 5

Strongly disagree Strongly agree

19. There is a chance to compete against other users. *

Mark only one oval.

1 2 3 4 5

Strongly disagree Strongly agree

20. There is a chance to collaborate with other users. *

Mark only one oval.

1 2 3 4 5

Strongly disagree Strongly agree

21. Users can utilise learned skills in practice. *

Mark only one oval.

1 2 3 4 5

Strongly disagree Strongly agree

Personal opinions.

In this section there are some open ended questions about how the program is generally perceived.

22. In your opinion, what are the greatest benefits of Duolingo compared to using other learning methods?

There are for example self-study books for learning languages, what makes Duolingo better.

.....

.....

.....

.....

.....

23. In your opinion, what are the greatest drawbacks of Duolingo compared to other learning methods.

.....

.....

.....

.....

.....

24. Have you used the skills learned in Duolingo in a real life situation? If so, how and where?

.....
.....
.....
.....

25. If you have used other self-learning tools similar to Duolingo, how would you compare them?

.....
.....
.....
.....

26. If you have any other opinions or comments about Duolingo, please feel free to express them here.

.....
.....
.....
.....

Background information.

Finally we would ask you to fill in some background information.

27. Nationality *

.....

28. Native language.

.....

29. Year of birth (YYYY) *

.....

30. Gender *

Mark only one oval.

- I'd rather not say
- Female
- Male

31. Highest Duolingo level in any language. *

.....

32. Do you use any other programs similar to Duolingo?

Mark only one oval.

- Yes
- No

33. If so, which ones.

You may also describe them briefly if you wish.

.....
.....
.....
.....

