University of Tartu

Faculty of Social Sciences

Institute of Education

Curriculum: Educational Technology

Joshua Schiefelbein

DEVELOPMENT AND USER EXPERIENCE EVALUATION OF

LANGUAGE LEARNING MOBILE APPLICATIONS

MA Thesis

Supervisor: Angeliki Eirini Chounta & Emanuele Bardone

Abstract (120 words)

In today's digital age, applications exist for everything from ordering pizza, applying to a university, and learning a new language. Furthermore, as the world becomes increasingly connected, knowing a second language is likely to make life easier. Using David Kolb's theory of the Experiential Learning Cycle as the key approach to studying a foreign language, this thesis studies the game design and gamification process undertaken during the development of two different mobile language learning applications before evaluating the user experience and user interface of low-fidelity paper prototypes created for both applications through a series of one-on-one interviews with individuals of different backgrounds.

Table of Contents

1 Introduction

As a foreign language instructor in Saint Petersburg, Russia, a common question asked to all students was "How did you study X language?" Usually, "X" language would be English, but it could have been others like German, Spanish, or French. Even in IELTS, the English language proficiency test that is commonly required for non-native English speaking students to pass in order to matriculate into an English-speaking program, one of the first prompts that test takers must speak on during the oral examination is "Tell us about your experience learning English as a foreign language." The answer for most Russian students is generally the same at first. Most students in school begin learning English from either first or second grade, and it is typically a required subject until the second year of university, which means many children study English for twelve years, yet several finish their English studies with a near C2 level of proficiency whereas others never advanced past A1.

Why? Russian education is defined by high levels of bureaucracy and centralization, so students often experience similar English programs across the country. One might suggest that the presence of foreigners in a city such as Moscow would mean that Russians in Moscow speak better than a place like Novouralsk, a closed city where no foreigners are permitted. Yet some students from Novouralsk can achieve near C2 and many in Moscow and Saint Petersburg find it difficult to string together enough words correctly to produce a basic sentence.

What is the difference in experiences between such students? How are their English skills so different despite having a near-identical curriculum? The answer may lie in gaming. Several students with high language levels reported that gaming, or rather online gaming with foreigners, oftentimes compelled them to use English, which provided them a significant amount of practice opportunities over their cohorts who may have either not been so inclined to game or they were more inclined to game with individuals who spoke the same language. Taking this phenomenon into consideration, gamification of language learning may prove to be an effective way of supporting students who learn a second language to engage more in the target language and thus learn through their experiences. To understand gamification, a concept derived from the word "game", one must explore what a game entails. In other terms, what makes a game a "game".

Jane McGonigal once wrote that games share four defining traits when all genre differences and technological complexities are removed: a goal, rules, a feedback system, and voluntary participation (McGonigal, 2011). According to her, the goal provides the sense of purpose in playing the game and it is the outcome players want to achieve. Rules are the limitations placed on how players can achieve the goal. By eliminating the easiest path to success, players are forced to think more creatively, critically, and strategically in finding a solution. The feedback system informs players how close they are to achieving their goal. The system can take many forms, such as points, levels, score, or progress bars. Real-time feedback reinforces trust by promising players that the journey to their goal is achievable, thereby keeping players motivated to continue. Lastly, games must indicate to all players that participation is voluntary. All users play the game knowingly and willingly, and that freedom to enter and exit a game at will ensures that the game's intentionally stressful and challenging requirements are experienced as a safe and pleasurable activity. McGonigal continues by saying that although many modern, advanced games boast high degrees of interactivity, narrative, virtual environments, and high-resolution graphics, these elements are not defining features. More important are the aforementioned four traits. All else is extra.

Papers, Please[1], a game released in 2013 about a border control official who must make decisions whether or not to stamp an individual's passport to permit entry into the fictional country of Arstotska, contains no advanced graphics. The game is rendered primarily 2-D with pixelated graphics, but because the four traits are achieved quite strongly through game mechanics that trigger an intense emotional reaction, the sense of immersion users feel while playing *Papers, Please* lifted the game to be considered one of the best released in 2013. Normally a game or film with the premise of passport control would not typically inspire great confidence for success. Yet, *Papers, Please* advanced the discussion of whether video games constitute a form of art. Lucas Pope, the creator of *Papers, Please*, eventually launched another game in 2018, *The Return of the Obra Dinn*[2], which had even lower resolution graphics than his previous one. *Return,* a tale about an insurance agent who inspects a Ghost Ship in the early 1800s to assess how the crew and passengers perished or disappeared, uses 1-bit graphics, the

---

[1] https://papersplea.se/
[2] https://obradinn.com/

kind used in early Macintosh computers, and an emphasis on the same four traits led it to receiving an even better critical review than *Papers*.

Even a game as basic as *Tetris*[3], a simple 2-D puzzle game featuring falling blocks that the player must stack them while leaving as few empty gaps as possible, remains highly popular since its release in 1984, ironic in light of the fact that *Tetris* is impossible to win. The longer a player engages the game, the harder it becomes as blocks become faster. Yet when analyzing the game's quality, it can be seen the four traits are effectively accomplished considering how the player receives a significant amount of real-time feedback that keeps engagement high. Among the different feedback mechanics are visual cues, such as how each completed row of blocks disappears instantly or a counter that displays an ever-increasing score, and emotional cues triggered by the rising degree of difficulty and challenge through the acceleration of blocks.

Huizinga (2000) describes games as non-serious and highly engaging voluntary activities establishing with a framework dictated by rules and social boundaries. Avedon et al (1971) advance the definition by indicating that conflict between equal parties is a necessity, and that the final result must be unequal. Salen et al's definition of a game declares that the game's end result should be a quantifiable outcome. Delving into the specifics of a game's constitution, Juul (2003) proposes that games include six primary features: rules, variables, quantifiable outcomes, value-laden outcomes, player investment, and negotiable consequences. Juul's six features, after some minor grouping and reclassification, largely reflect McGonigal's assertion that games share four defining traits.

Due to the popularity of games, the prevalence of mobile devices such as smartphones and tablets, and size of the gaming industry, many companies are strongly incentivized to gamify their products. According to statistics gathered by WePC and Batchelor (2019), by 2020, the gaming market is expected to be valued at over 90 billion U.S. dollars after generating $135 billion in revenue in 2018, a 10.9 percent increase from 2017. Software developers are also attentive towards the growing industry as 38 percent of game developers work on mobile games since 47 percent of 2018's revenue came from mobile devices.

---

[3] https://tetris.com/

One industry that is making significant efforts to gamify and monetize is the language learning industry. For the better part of a decade, Rosetta Stone, a computer-assisted language learning software, was credited the preeminent example of language learning applications due to its role as applications designed to help individuals learn a foreign language. However, Rosetta Stone is criticized for a weak and sometimes culturally irrelevant pedagogical approach and, perhaps more importantly, a price tag of over 124 dollars to obtain the necessary user license for one level, a prohibitive cost for the casual user. Capitalizing on the dearth in language learning software and the ease of downloading applications to a smartphone, the number of mobile language learning apps has skyrocketed. The top applications share a defining philosophy of using elements of gamification to aid user retention and improve engagement. A drawback to some apps, however, is their monetization strategy, sometimes seen as overly aggressive or exploitive, which relies on ad-clicks or *freemium* models. Freemium, a combination of the words *free* and *premium,* indicates that the app is free to download but only limited content. To unlock all content, users must typically pay either a one-time fee or purchase a subscription.

Language learning applications exist for a variety of gamers. According to gaming jargon, most gamers fall into the categories hardcore, tryhard, retro, professional, casual, or hypercasual, the final one describing gamers who only play for five minutes before stopping. In the 1990s, Richard Bartle described four types of players: Killers, Explorers, Achievers, and Socializers. Bartle's framework largely reflects gaming slang in that hardcore gamers are Killers, tryhards are Explorers, professionals are Achievers, and retros/casuals/hypercasuals are Socializers. Taking into account gaming behaviors is an essential user design consideration when developing a game. As an example, for hypercasual users, their preferred language learning application might likely be Drops, which allows users to build their vocabulary by learning nouns. Drops is designed for individuals who only wish to commit five minutes or so per day towards learning the target language. As such, the content of Drops focuses primarily on nouns and does not provide much linguistic support in grammar, usage, conjugations, or speaking. Due to its short noun vocabulary tasks, it is considered an ideal hypercasual app. But for language learners with a more long-term perspective and the willingness to invest time, other apps such as DuoLingo are more popular and potentially more effective.

Arguably the current market leader in language learning apps, DuoLingo features a holistic approach to studying as users engage in vocabulary, grammar, and usage simultaneously with lessons of ten questions that employ illustrated flash cards and fill-in-the-blank exercises as challenges. Since its first launch, DuoLingo has continuously updated its capabilities by integrating voice recognition so that users can practice pronunciation and audio files to facilitate listening comprehension. Tasks are varied as users sometimes have to tap buttons to select answers, speak into the microphone, drag tiles into the correct sequence, or type. Gamification has been strongly integrated as users receive instant in-lesson feedback on whether answers are correct or wrong, streak counts tally how many days individuals used the app consecutively, levels that indicate the degree of difficulty achieved within a topic, and a health counter that tracks how many potential questions a user can answer incorrectly before rendered unable to continue until the health counter refreshes. DuoLingo's main avatar, an owl named "Duo", is used to establish an emotional connection with the user, and if individuals have not accessed the app for some period of time, notifications are generated using "Duo" as an emotional reason for the individual to log into the app. For achievers or professional gamers who wish to display their accomplishments, awards are included such as merits, badges, prizes, and in-game currency rewards that can be cashed to acquire access to specialized linguistic content, such as how to flirt in the target language. Socializers can connect with their friends who are also using the app, and Killers can compete with friends to see who has studied the most or progressed the farthest. Pedagogically speaking, material is scaffolded when questions answered incorrectly are asked again before the lesson can be considered completed. However, the progression of target vocabulary and the overall curricular approach is not always clear to ordinary users. Pop culture enthusiasts can even study languages like Esperanto, Klingon (from Star Trek), and High Valyrian (from Game of Thrones).

Memrise, one of DuoLingo's primary competitors, has several capabilities such as supporting offline courses and numerous languages. However, in contrast to DuoLingo which begins its courses with common words in the language like boy and girl, Memrise begins by pairing sounds and sentences in the target language with similar sounds and sentences to the medium language. Images of the word and videos of speakers reinforce word association, while

gamification is employing through the inclusion of levels, setting daily goals, scoring lessons not only by the number of correct answers but also speed and streak bonuses, pressure from a countdown timer, and an animation of a seed growing into a flower. A community element exists where socializers can also not only connect with friends, but they can follow other users. Whereas DuoLingo has no strong narrative element, Memrise has the user become a spy that navigates through different worlds, earning points for correct answers. Similar to DuoLingo's special content like flirting, Memrise includes vocabulary collections such as Japanese naughty words or excerpts from Harry Potter and the Sorcerer's Stone, in addition to the traditional textbook-style lists. Ultimately, its core learning approach relies on humorous or bizarre associations through spaced repetition and mnemonics, the former being determined by a special algorithm that calculates the frequency the user should review material.

In contrast to the above two market leaders in language learning apps, Mondly is more reminiscent of traditional classroom-based instruction regarding its comprehensive and structured approach. The basic package provides users an initial start in the target language, and grammar charts on topics like conjugation and declension provide studious users an opportunity to review grammar, an element mostly missing from DuoLingo and Memrise thereby making Mondly's capability to provide grammatical references a standout feature. A simple tap on the verb will display its conjugations and translation. However, the majority of the content is locked to the average user and requires a subscription, and the interface is regarded as less straightforward than its two key rivals. Mondly does include speech recognition for pronunciation activities, gamification features such as level ups and leaderboards, and translation drills put the user's skills to the test.

Another application, one not as widespread as Drops, DuoLingo, Memrise, or Mondly, is Lingvist. The story of its founder, Mait Muntel, fits the user scenario of nearly every individual who tries to use an app to learn a language. He lived for many years in a part of a country where he did not speak the local language, but due to his work, he did not have time to attend classes while available digital tools and language failed to provide the vocabulary he wanted. Consequently, in launching Lingvist, the company attempts to accelerate language learning up to ten times faster through using big data and artificial intelligence, the sum of which maps the

knowledge and skills of each learn to adapt learning material in real time and present a unique yet challenging experience. Lingvist's minimalist design where nothing more than necessary is included, such as just enough analytics to track and monitor progress, and relatively small amount of gamification has led to it being considered a useful companion app to one of the above gamified apps. The app has even garnered a significant amount of attention on DuoLingo's community forums.

1.1 Purpose of Research

This thesis will look into the effects of gamification on the user's learning experience of a mobile application aimed at improving a user's English skills and whether gamification is a key element of success in a language learning application and if gamification should take priority over content. Should gamification be indicated as a key element of success, then such a conclusion might imply that applications such as DuoLingo have adopted the correct approach, and therefore future language learning applications ought to employ gamification strategies. Consequently, this paper has three research questions. Does gamification improve the learning experience of language learners when using an application? Is gamification more effective for younger audiences as they are less inclined to focus on grammar materials and would rather immerse themselves into an environment? Are language applications best used in pairs where one application focuses more on grammar and vocabulary learning and retention while the second application concentrates on practice and spontaneous reproduction?

To that end, this paper recounts the design of two original applications two applications - LearnIT ASAP and Starfighter - for learning English as a second language using two different pedagogical approaches. LearnIT ASAP was designed following a traditional approach to language learning while Starfighter implemented several gamification elements. Afterwards, extensive, in-depth interviews were conducted with eleven individuals from different backgrounds who were tasked to interact with the applications' prototypes. The purpose of these interviews was to record the users' experience with respect to the applications' effectiveness at achieving the purpose of learning the target language and why. The results were transcribed,

analyzed, and coded to aid the comparison of approaches, which ultimately provided answers to the research questions.

The findings suggest that gamification can support long-term user retention, but the gamified applications must also provide some degree of traditional instruction, with respect to grammatical and vocabulary references, to help guide users towards language proficiency. The ideal approach in this regards would be one of two options. Either an application that combines traditional learning with a built-in gaming environment or two applications to be used as companion applications where one provides the theoretical language background while the other provides the ability to practice.

2 Literature Review

2.1 Philosophies of Learning

When it comes to education and theories of learning, the three basic types of learning theory are behaviorist, cognitive constructivist, and social constructivist. Behaviorism posits that learning is the result of behaviors acquired through conditioning in that individuals' responses to environmental stimuli shape actions (Watson, 1913). The conclusion is that any individual can be trained to perform a task regardless of genetic background, personality, and cognition, so long as it is within physical capabilities. From these three schools, many different approaches have been discussed, defined, and researched. Cognitive constructivism argues that knowledge is a network of active systems generated by previous learning experiences such that every learner then processes new experiences and information using the network, which is shaped by their present knowledge, stage of cognitive development, cultural background, and many other personal factors. Effectively, learning is an active process of construction where learners build the world they perceive around them as opposed to passively absorbing knowledge like in behaviorism (Piaget, 1936). Social constructivism advances the theory of constructivism a step farther by emphasizing the prominence of language and culture in learning in helping learners construct their reality where language helps scaffold learners to a higher level of knowledge, thus

allowing them to achieve their potential development (Vygotsky, 1978). Some organizations create a spinoff theory that bases itself in a hybrid of the three main theories. Moodle[4], for example, states that its pedagogical philosophy uses social constructionism, which is the belief that groups construct knowledge from one another in a social environment with shared artifacts and shared meanings.

For the language apps being designed, the primary learning theory guiding their development is experiential learning, which was greatly developed and influenced by David Kolb, who drew on the work of John Dewey's pragmatism and Jean Piaget's constructivism (Kolb, 1975). Pragmatism holds that individuals learn best by being hands-on and practical, and that education should focus on real-world applications of learned material (Dewey, 1916). From Dewey's standpoint, students can only learn by interacting and adapting to their environment. Piaget's constructivism is similar to Dewey's pragmatism in that constructivists argue learning occurs when people produce knowledge and form meaning based on their experiences (Piaget, 1936).

Experiential learning essentially combines the essence of these two philosophies into one by stating that learning occurs through the intentional reflection on doing some process. An important difference between pragmatism's hands-on learning and experiential learning is that hands-on learning does not necessitate reflection, whereas experiential learning does. To illustrate his point, Kolb crafted the Experiential Learning Cycle, which is composed of four steps: concrete experience, observation and reflection on that experience, formation of abstract concepts based on the reflection, and active experimentation with the new concept. The cycle repeats continuously with no stopping so long as the learner continues trying to gain knowledge, which can be taken through both personal and environmental experiences, and the cycle may begin at any one of the four elements, but usually it starts with a concrete experience. However, Kolb states that in order to acquire genuine knowledge from an experience, the learner must have four abilities: willingness to be actively involved in the experience; capable of reflecting on the experience; possessing and using analytical skills to conceptualize the experience; and possess decision making and problem solving skills to use the new ideas taken from the experience.

---

[4] https://docs.moodle.org/37/en/Philosophy

With regards to language learning approaches, numerous methods exist (Richards et al). Some of the more popular methods can be grouped by categories. For example, approaches where all communication is conducted in the target language include the Direct Method (almost no grammar but with an emphasis on good pronunciation), Communicative Language Teaching (situational speaking), the Silent Way (the teacher says as little as possible, forcing the learner to speak), Immersion, and Task-based language learning (students use the language they possess to complete the set task). A second category is more behavioralist, such as Total Physical Response (completing instructions) and Audio-Lingual (learning through acquiring habits). The third category is more constructivist, including the Lexical Syllabus (computer analysis of most common words), Grammar-Translation, and the Structural Approach (language is a set of grammar rules to be learned).

Language learning apps typically combine a mix of approaches from the first category and the third category, depending on the specific approach the app wishes to employ. However, to retain users, apps need some mechanism, for which they have typically resorted to gamification.

2.2 Gamification

Gamification is defined as the usage of game-play mechanics in a non-game context rather than a fully developed game, oftentimes to improve the user experience and user engagement in non-game services and applications (Deterding et al, 2011). It involves applying elements of *gamefulness, gameful interaction,* and *gameful design* with a specific intention in mind, gamefulness is the lived experience, gameful interaction is the objects, tools, and contexts that bring about gamefulness, and gameful design is the practice of crafting gamefulness. Emphasizing that gamification is a process of enhancing services with motivational affordances, Deterding insists those affordances must be the same as the ones used in games regardless of outcome. Based on this conceptualization, gamification has three main parts: implemented motivational affordances, resulting psychological outcomes, and further behavioral outcomes. Consequently, with this conceptualization, many different everyday occurrences can be

gamified, such as applications, tasks, processes, and even contexts themselves. Coined by Terrill (2008), 'gameification' was "taking game mechanics and applying them to other web properties to increase engagement." The definition and purpose of gamification has been slightly expanded and made more specific as increasing user engagement through the application of game-like mechanics such as scoreboards and rapid personalized feedback (Flatla et al, 2011). This increase in user engagement correlates with an increase in reported personal ownership (Pavlus 2010). It is heavily used in non-game web and mobile applications such as DuoLingo, Lingvist, and Habitica, oftentimes in order to convince individuals to download the application and retain users over a long term. Gamification can also be integrated into existing companies and services, such as Codecademy, which uses game-like elements to help teach users to code), or Badgeville, which companies hire to use employee to improve and reward reams and programs in the hopes of maximizing long-term growth. Chrons and Sundekk (2011) explain that the combination of mundane activities and gamification, individuals are more likely to continue using the application, especially because obligatory mundane activities such as cleaning the house are not appealing.

Huotari and Hamari (2012) argue that Deterding's definition is too limited and focuses too must on the systematic approach of gamification. Incorporating the experiential value of gaming, they offer the definition of gamification as "a process of enhancing a service with affordance for gameful experiences in order to support user's overall value creation." In place of Deterding's structural perspective and concentration on the methods, Huotari and Hamari focus on the goal of gamification and the experiences that are targeted, elaborating that gamifiers attempt to increase the likelihood of gameful experiences by imbuing the service with motivational affordances. To them, simply the inclusion of gaming elements does not automatically create a gameful experience, nor does the process of gamification need to be successful. More important is that gamification is able to support users in creating gameful experiences.

Gamification also implies a certain quality of social interaction where the user engages with other players. When users perceive some social presence, they are more likely to respond and demonstrate feelings like empathy or anger, or follow norms of social conduct like taking

turns (Fogg, 2002). Fogg also explores human-computer interaction by proposing the Fogg's Behavior Model (FBM) (Fogg, 2009), which studies factors and how they cause certain behaviors. His model consists of three primary elements: motivation (as Deterding might describe as motivational affordance), ability, and triggers. If those three elements happen at the same time, it can determine a target behavior. Applied to education, a student must be simultaneously motivated, capable, and triggered to accomplish the target objective. To achieve this, the student must be in a state of 'flow', which is essentially concentration.

Motivation can be designed by the application of opposite emotions like happiness and sadness, hope and fear, and acceptance or rejection. Even if a student proves capable of solving a problem, the student will not do so if no motivation exists. Once when there is a positive or negative effect, such as a high grade or damaged reputation, might the student do the task.

As for ability, even highly motivated students cannot solve a task if they lack the necessary skill. Language learning is a good example of this. Many individuals may have a strong desire to visit Italy and speak Italian with locals, but without any practice or training, try as they might, those individuals will prove incapable of maintaining a conversation in Italian.

Fogg indicates that motivation and ability are not enough to predict behaviors. To do this, a target behavior must have a trigger, in other words a call to action. There must be something that instructs the student to complete a specific action at a given time.

Muntean (2011) concluded that gamification does not entail creating a game, but rather it is a means to make education more fun and engaging without undermining its credibility. When gamification occurs and game mechanics and features are employed, each mechanic is defined by three attributes: game mechanic type or motivational affordance (progression, feedback, behavioral), benefits (loyalty, engagement, time spent, influence, fun), and personality types (explorer, socializer, killer, and achiever).

In an empirical survey of the different studies conducted on gamification, Hamari et al. (2014) identify 10 categories of motivational affordance examined in previous literature: points, leaderboards, achievements/badges, levels, story/theme, clear goals, feedback, rewards, progress, and challenges. Of these ten categories, the most commonly found affordances used are points, leaderboards, and badges. An experimental study by Sailer et al (2017) concluded that

gamification itself was not necessarily effective, but different elements possessed some affect. Sailer indicated that badges, leaderboards, and performance graphs positively impacted competence need satisfaction and task meaningfulness, while avatars, meaningful stories, and teammates enhanced social relatedness. Farzan et al (2008) indicates that the effects of gamification may be more of a short-term benefit than long-term due to the novelty effect, research by Thom et al (2012) asserts that detrimental effects may occur to users who remain engaged by gamification if the motivational affordances are removed.

According to Hamari et al (2014), most studies concluded that the learning outcomes of gamification were mostly positive as they led to increased motivation, engagement, and enjoyment. However, some of the studies analyzed discuss potential negative consequences to be watchful for, such as fallout of increased competition, task evaluation difficulties, and design features.

In designing meaningful gamification (Nicholson 2015), defined as the use of gameful and playful layers to help a user find personal connections that motivate engagement with a specific context for long-term change, Nicholson explains that reward-based gamification is useful for short-term goals and situations where the participants have no personal connections or intrinsic motivation to engage in a context, rewards can reduce intrinsic motivation and the long-term desire to engage with the real world context. If the goal is long-term change, then rewards should be avoided and other game-based elements used to create a system based on concepts of meaningful gamification. Nicholson proposed that six elements, inspired by game design, must be considered in designing the gamification: Play (the freedom to explore and fail within the environment), Exposition (the stories for participants before allowing participants to craft their own), Choice (developing systems allow participants to make decisions), Information (the usage of game design and game display to allow participants to explore), Engagement (the encouraging of participants to discover and learn from others), and Reflection (the assisting of participants to find other interests and experiences to enhance engagement and learning). Rearranged, the letters spell RECIPE.

Zichermann and Linder explain that gamification is a way to supplement branding initiatives (2010). By evaluating motivation in relation to psychology, Zichermann dissects

motivation into two types: intrinsic, where activities occur because they are in accordance with an individual's values, and extrinsic, when external rewards such as money or status are given as a result of taking part in an activity. In his opinion, designers should craft extrinsic motivators so that they feel more intrinsic as purely intrinsic motivators are variable and unreliable. However, Zichermann (2011) claims that financial reward, a traditional extrinsic motivator, simultaneously decreases motivation while improving performance. To balance this effect, monetary and generic non-monetary incentives need to be offered.

Games, like McGonigal stated, need to permit voluntary participation, which means users need to be allowed some degree of autonomy, competence evaluation, and social relatedness. Aparicio et al (2012) identified different game elements that support those three traits. Game elements such as profile pages, avatars, customizable interfaces, and privacy/notification controls give users a sense of autonomy. Competence evaluation is achieved through positive feedback, optimized challenged, points, levels, leaderboards, and intuitive controls. Functionalities like groups, chats, blogs, and social network integration support social relatedness.

Much of the above literature on gamification focuses on the implementation and gamification of activities, as well as the purpose behind it, but less of it focuses on evaluating the user's experience. Compared to articles written on game design and gamification, there is a shortage of articles that measure the user experience after the gamification of some task.

2.3 User Experience Evaluation

Hassenzahl (2008) provides a two-part definition of user experience: 1) "a momentary, primarily evaluative feeling (good-bad) while interacting with a product or service"; 2) "Good UX is the consequence of fulfilling the human needs for autonomy, competency, stimulation (self-oriented), relatedness, and popularity (others-oriented) through interacting with the product or service (i.e., hedonic quality). Pragamatic quality facilitates the potential fulfilment of be-goals." Hassenzahl explains that people perceive interactive products in two ways: pragmatic quality and hedonic quality. Pragmatic quality is the product's perceived ability to facilitate a user to achieve some task, such as ordering a book or making a phone call. In other words, pragmatic quality is the product's utility and usability with regards to a task. Hedonic quality is

the product's perceived ability to scaffold a user's capability to achieve some competence, which he describes as "be-goals." Jakob Nielsen's ten usability heuristics provide the initial guide to coding user experience interviews (Nielsen).

Through a mixture of Anticipated eXperience Evaluations (AXE) and Contextual Laddering, qualitative data can be gathered during user experience interviews. AXE provides an initial perspective on the user experience of a product by using visual stimuli to encourage interviewees to discuss their attitudes, practices, and assessments. The AXE approach requires three distinct steps in the interview: concept briefing (including the use of mock-ups or paper prototypes), concept evaluation, and data analysis (including coding of data by a list of features, such as Nielsen's heuristics). Contextual Laddering requires the interviewers to explore the reasons why certain elements of a product are liked or disliked, thus necessitating several "why" questions.

3 Methodology

3.1 Pedagogical Approach

3.1.1 LearnIT ASAP

Although the learning philosophy behind LearnIT ASAP is Kolb's experiential learning, the pedagogical approach is more traditional in its constructivist approach, primarily due to the reliance of using fill-in-the-blank questions that can be scaffolded through grammar review. This approach is primarily because of the prominence of the "Use of English" section that is very popular in Cambridge English Exams in indicating an individual's English ability according to the Common European Framework of Reference for Languages. Cambridge tests, which include the Key English Test (KET for A2), Preliminary English Test (PET for B1), First Certificate in English (FCE for B2), Certificate of Advanced English (CAE for C1), and Certificate of Proficiency in English (CPE for C2), all include a reading section that contains one, two, or three parts that evaluate a student's use of English skills. Use of English can entail selecting multiple choice questions, fill in the blank, or word formation, the latter being where a test taker receives

a base verb and must decline the verb correctly into a noun, adjective, or adverb depending on the context of the question.

3.1.2 Starfighter

Starfighter relies more heavily on gamification and experiential learning as its guiding paradigms, while the language learning approach is more akin to methods of drilling, such as the one employed in the Rassias Method, which seeks to accelerate the language learning process through controlled practice that will ultimately lead to language retention and spontaneous use of target material. A similar strategy is employed in DuoLingo as questions answered incorrectly are skipped and then returned to at the end, compelling the user to attempt the question once more in order to complete the level.

3.2 Content

In both apps, the content was written in JavaScript Object Notation (JSON). However, both apps were annotated in a slightly different style primarily because the primary interaction pattern was different between them. The value of using JSON is that the script is a simple way to read text into a game. It stores data in text form, it's lightweight, and it allows for easy transfer and data storage.

3.2.1 LearnIT ASAP

In LearnIT ASAP, the content is comprised of several short funny anecdotes where certain words are eliminated that the user must then fill in. Areas of grammar and vocabulary include general topics like prepositions, collocations, conjugations, and other basic elements of grammar. No images, audio, or video supplementary materials are provided.

For LearnIT ASAP's JSON files, an asterisk (*) was written into the text in place of a word, and corresponding to the asterisk, an array with possible options was created. The asterisk is used to indicate there will be a dropdown menu. Answer options are then inserted into the

array with the boolean true/false where true indicates the correct response and false the incorrect response.

3.2.2 Starfighter

In Starfighter, there is a default path of 200 questions covering a wide variety of grammar and vocabulary topics, and then additional tests that target a specific area of grammar or vocabulary is made available for 20 questions total. Areas of grammar and vocabulary include general topics like prepositions, collocations, conjugations, and other basic elements of grammar. The questions within a topic focus on frequent mistakes made by non-native speakers. This might include phrasal verbs like "pick up" versus "pick out" or articles such as "the Czech Republic" versus "Germany". No images, audio, or video supplementary materials are provided, nor is context.

The JSON script is easier to write and understand that LearnIT ASAP. Each question includes the blank written in underscores (_) instead of an asterisk. Then an array is constructed with two or three possible options, depending on the topic.

## 3.3 Interface Design

3.3.1 LearnIT ASAP
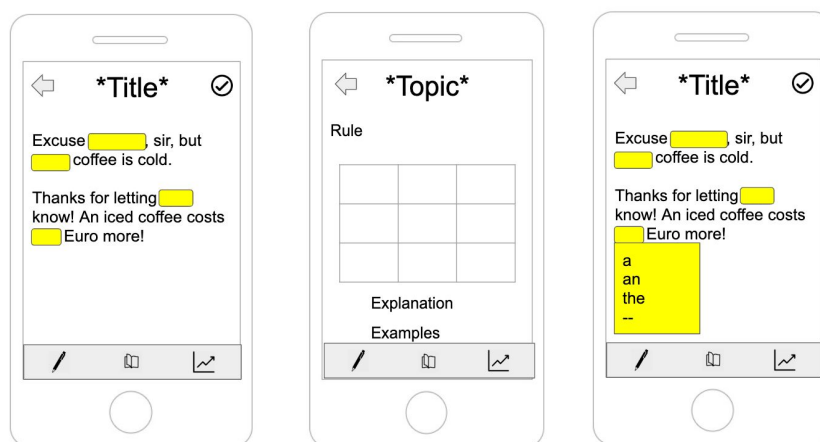
Figure 1. Example screens from LearnIT ASAP showing the task completion system.

The interface of LearnIT ASAP strived to be simplistic using a website paradigm of interface with minimal text considering the content of the app is primarily text based. On the main menu, three simple buttons exist: Start, Tests, and Grammar. To navigate to each section, the button should be tapped. To exit the app completely, the main screen button of the phone can be pressed.

By tapping on Start, the user is directed to the first task, as seen in the first screenshot in Figure 1. Words are removed and replaced by yellow fields that represent buttons. Each test in LearnIT ASAP includes between three and five yellow buttons that must all be answered before the icon in the top-right corner can be tapped. A dropdown menu with four possible answers is revealed by tapping on a yellow button, as seen in the third screen in Figure 1, and to select an answer, the user need only tap on the answer of choice. Once the user has filled in all the yellow fields with a possible answer, the user should tap on the check mark in the circle at the top right of the screen, as seen in the first and third screenshots in Figure 1. The grammar page includes essential information regarding the topic of grammar being practiced in the current exercise, as seen in the second screenshot in Figure 1. If the exercise is on first conditional, the grammar page will display the rule of first conditional. If the exercise is on prepositions of time, then the grammar page will show the rules of prepositions of time, and so on.

The bottom menu bar, which can be seen in all three screenshots in Figure 1, is a quick navigation bar with three icons: a pen, a book, and a graph. The pen represents the exercise, the book means grammar, and the graph will show the user's statistical progress. Statistics track the number of tests available, the number of tests unlocked, the number of tests completed, and the average time taken to complete a test. The quick navigation bar allows the user the chance to avoid tapping the back arrow, located in the top left corner, in order to return to the previous screen.
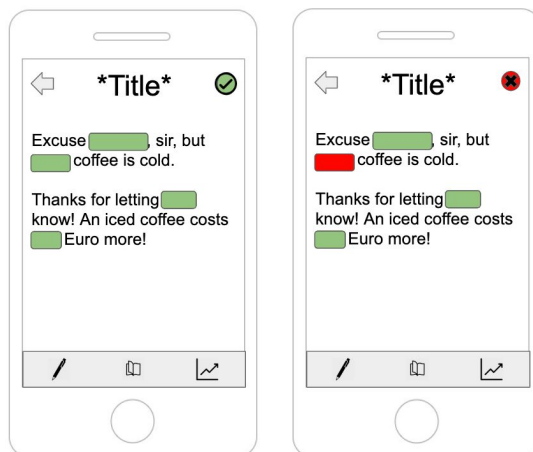
Figure 2. Example screens from LearnIT ASAP showing the feedback system.

Once the user submitted the possible answers, the system will grade the submitted answers and display which are correct and incorrect by coloring the yellow fields as green for correct answers and red for incorrect answers, as seen in both screenshots in Figure 2. Should all fields be correct, then the check mark circle in the top right becomes green, as seen in the first screenshot in Figure 2. To continue on to the next exercise, the user should tap on the green check mark circle to advance. The rationale behind this design choice is to provide the user a final opportunity to quickly review the grammar and understand why the answers were graded as correct if the user so desires.

If at least one of the fields is incorrect and colored red, the check mark circle will become red and the check mark will be replaced by an X, as seen in the second screenshot in Figure 2. Since the user must have all fields colored red, the user must resubmit the answers. To accomplish this, the user needs to tap again on the red fields and select a different response. However, to avoid gaming the system by using brute force problem solving, the possible responses will have been shuffled into a different order to compel the user to think critically or refer to the grammar page in determining the correct response. Failure to think critically about the answer puts the learner in risk of an inconvenient loop of wrong answers. Once a different response is selected, the red field disappears, as does the red circle with the black X, which is replaced by the original white circle with the checkmark. Once the user is satisfied with the new responses, the user resubmits by tapping the checkmarked circle as before.
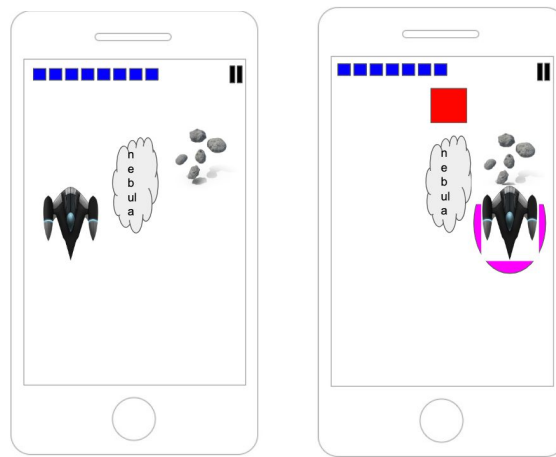
3.3.2 Starfighter
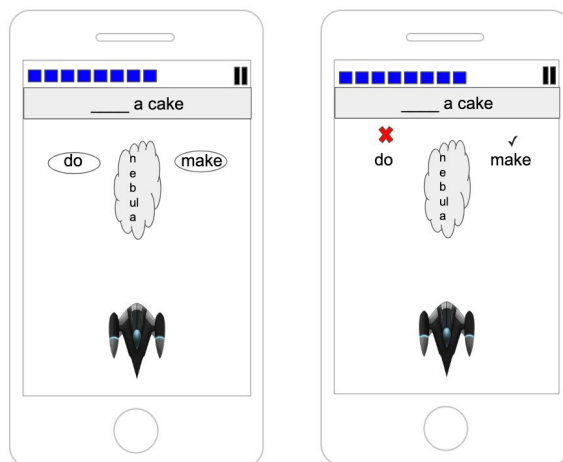
Figure 3. Example screens from Starfighter. The first screen shows a
typical question, while the second shows how feedback is provided in game.

In much the same style, Starfighter, which uses a gaming paradigm of interface, is fairly minimalistic outside of gameplay. The main menu includes three tappable options that function like buttons: Start, Skill Check, and Leaderboard. By tapping on Start, the user automatically begins gameplay, which can be seen in both screenshots in Figure 3. Skill Check allows the user to select specific target grammar and vocabulary areas such as prepositions or antonyms to practice in sets of twenty questions. A scroll button is included in the form of an arrow pointing downwards, and the user can either tap the button or drag a finger in order to scroll through the different options available to practice. Leaderboard tracks the scores of app users according to two categories: Overall and Friends. Overall displays the top scores of all app users in relation to the user's personal score, while Friends displays the scores of friends that the user is connected to again in relation to the user's personal score. A back button is placed at the bottom of each screen to allow the user to navigate back to the previous screen. To avoid having to hit the back button repeatedly, most screens are available within two taps of the main menu.

In gameplay, the primary user interface, seen in the first screenshot in Figure 3, is arranged with the ship at the bottom of the screen against a space background filled with stars, a pause button in the top right corner for when the user may wish to temporarily halt gameplay, a shield hit point counter in the top left corner characterized by eight squares, a stripe near the top

with a text overlay communicating the question, two buttons that contain possible answers, and a nebula between the answer buttons. From the moment the user begins playing, the ship is in constant motion, but similar to arcade games like *Galaga* or *Galaxian*, the ship remains at the bottom of the screen as the nebula approaches the ship. The answer buttons remain stationary, and the user must tap one button or the other to avoid striking the nebula and losing some of the shield's strength. The user is then shown which answer is correct and which is wrong, as seen in the second screenshot in Figure 3.

Figure 4. Example screens from Starfighter showing possible animation cutscenes.

After the user receives feedback about the correct answer, the starfighter will veer to the left or right depending on the answer the user chose, the process of which can be seen in both screenshots in Figure 4. If the user selected the correct answer, the ship is shown flying to the side that does not have meteorites in its path, an example of which is seen in the first screenshot in Figure 4. If the user selected the wrong answer, the ship is shown flying directly into the meteorites, an example of which is seen in the second screenshot in Figure 4. This causes the shields to flash and a red square appears in the top center of the screen to signify the shield's strength has dropped, leaving $n - 1$ blue squares for the current shield strength. To avoid gaming the system, the questions will be randomized and the answer buttons will be randomized in terms of their placement on the left or right side.

3.4 Gamification

Gamification Mechanics	LearnIT ASAP	Starfighter
Points		x
Badges		
Leaderboards		x
Relationships		x
Achievements/Challenges		
Constraints	x	x
Narrative		x
Feedback	x	x
Journey		x
Emotion	x	x
Currency		
Prizes/Awards/Bonuses		x

Table 1. A side-by-side comparison displaying gamification mechanics in both applications.

3.4.1 LearnIT ASAP

From a gamification perspective, LearnIT ASAP uses little to no gamification elements as it has a more traditional approach to language learning. The primary elements of gamification it has is the need to complete all steps before moving to the next level, tracking personal user statistics in the form of the average amount of time taken to complete a task, and clear, instant feedback upon submission. Feedback is necessary to help engage the user and guide the learning process, while the other two provide benchmarkers for the user to self-evaluate progress. Other elements such as leaderboards, prizes, awards, achievements, daily challenges, or other popular gaming mechanisms do not exist.

3.4.2 Starfighter

Starfighter includes several gaming mechanisms implemented. First, there is a clear points counter. Users receive points based on how many questions they answer correctly, and then a multiplier effect exists depending on how many were answered correctly consecutively. Furthermore, the points counter allows users to compare their progress with a broader community through the usage of Leaderboards that track both overall app users and friend users. This sense of community and competition typically helps engage and motivate users to play the app more frequently, as seen in DuoLingo.

Secondly, there is a sense of pressure. The more questions a user answers, the faster the user must answer the next time, similar to the mechanic in *Tetris* when blocks fall faster the longer the user plays. This sense of pressure not only engages the user through the increasing challenge of gameplay, but it gives the user a sense of progress. The deployment of shields helps insulate users to some extent by protecting them from failing too quickly as speed rapidly increases, but the sense of dwindling shields can increase the pressure experienced in conjunction with time pressure.

Thirdly, checkpoints are used in the default path of questions. Since the default path consists of 200 questions, it is unreasonable to assume a player will attempt to play through all 200 questions in one sitting. Instead, the user is given a checkpoint update at every 20th question. If the user survived the 20 questions, then the user's shields will recharge, but if the user received between 12-14 correct responses, then the user would have to repeat the set of 20 questions before advancing past the checkpoint. The updates are critical in giving the user consistent feedback, which is also provided through points, increasing speed, shield hit points counter, and indication of right versus wrong answers.

Originally, the intent for Starfighter was not to have a starfighter navigating through a field of meteorites and nebulas, but rather the starfighter was supposed to fly through an asteroid field and destroy asteroids. If the correct answer was selected, then the starfighter would shoot and destroy the asteroid well enough to pass through the meteorites with no damage and the rubble would harmlessly bounce off the shields. If the wrong answer was selected, then the starfighter would only partially damage the asteroid, thereby leaving large enough chunks of rock that would partially drain the ship's shields. However, after discussion with game designers,

the designers highlighted that logically it made little sense why the ship's lasers would destroy one asteroid successfully and not destroy the other. Moreover, they pointed out that if it were a true logical game, the ship could simply try to fly around the asteroid. As a result, based on the acquired feedback, the focus of the game shifted away from destroying asteroids and instead went to safely navigating through a nebula field.

3.5 Approach

Interviews took place on a one-on-one basis over a two-week period at a location deemed convenient for both parties. Most interviews occurred on a weekday in the afternoon in a university office space, while two interviews were conducted at a coffee shop during relatively low traffic hours. The average duration for a single interview was approximately 35 minutes, though a few lasted almost up to an hour. Data was collected both in written form and in video, except for one interviewee who opted out of having the interview recorded. For that interview, all data was logged in written form.

3.5.1 Set-up

Eleven individuals were identified to take part in interviews that evaluated the usability and navigability of both applications. Participants of the final version of the protocol script came from a diverse array of backgrounds and ages with the youngest being 20 years old and the oldest being 50 years old. The average age of the participants was 29.5 with the median at 28. The majority of participants fell within the 18-29 age range with eight individuals. Only one participant was a Native English Speaker, while the other ten had English levels that were either B2 Upper Intermediate or higher. All interviewees spoke at least one foreign language, and most spoke at least two or three with English the most common followed by other Indo-European languages like German, French, Spanish, and Russia. The most represented group is Estonian with two native Estonians participating. Other represented nationalities include Colombia, Iran, Jamaica, Germany, India, Russia, America, Morocco, and Italy.

All participants were enrolled in or had completed some level of post-secondary education. Four interviewees were currently enrolled in a bachelor's program, while two were currently enrolled in a master's program. However, for the two currently enrolled in a master's program, both of them had completed a first master's degree and were now on their second. Three participants were current Ph.D. students and two had completely finished their Ph.D. program and are currently working as university researchers. Of the eleven participants, eight had studied at a foreign university as either a full-time or exchange student. Only one participant had studied a field not related to technology, engineering, or computer science. With the exception of two participants, all other participants reported having a job or professional background related to education or academic research. The remaining two participants reported working in different capacities for an information technology company as either a project manager or editor.

Regarding language applications, seven reported having used such an application at some point prior to the interview. Eight individuals indicated that learners who wish to use a mobile language application may find the process effective, but three indicated uncertainty regarding the level of effectiveness. Of the individuals who self-reported having used a language application, DuoLingo was the most commonly used with five participants indicating use, while Memrise had two, Babbel one, Drops one, and Keeleklikk one. Some individuals who used an application were unable to remember the name of the application they had used.

3.5.2 User Experience Interviews Protocol

Prior to conducting user experience interviews, an interview protocol needed to be generated. The interview protocol included a basic introduction that informed the interviewee of his or her basic rights, explained that there were no right or wrong answers, prompted the interviewee to review and sign a consent form agreeing to be recorded, and offered a chance for final questions of clarification about the upcoming interview.

Once the draft interview protocol was finished, a test session to evaluate the protocol and app screens was conducted. One individual was used as a trial subject to test the initial

wireframes and protocol script. After finishing the trial interview, the script was superficially edited for style, simplicity, and concision. The trial interview also provided a chance to test the printed wireframes to learn if misprints occurred or screens were missing.

The final interview protocol included two user scenarios, one for each application along with a one sentence tagline to generally summarize the primary mechanic. User scenarios were deemed necessary in order to better contextualize the interviewee, thereby achieving the contextualized laddering approach of user experience interviews, as to the purpose of the application and its target audience so that as they interacted with the application, they would be perceiving the application from the perspective of the target user.

The tagline for LearnIT ASAP was presented as a language learning app that provides individuals with a short story with several blanks that must be filled in correctly before the user can progress to the next story. The user scenario of LearnIT ASAP tells the story of Alexandra, a university-age Russian student with a B1 level of English. She has basic skills in English and can read basic texts and understand, but she really wants to improve her skills. She enjoys immersing herself in grammar and trying to follow the rules, but sometimes it is really hard for her to determine which words or word forms she needs to use when speaking or writing. Moreover, in the near future, she wants to take an English proficiency exam for a foreign study program, so she wants to use a tool that practices a similar skill as the one in exams, but with potentially more interesting or funny material. Also crucial is that the tool provide information about why she is or isn't right in answering a question.

This user scenario was created in this fashion as it touches on the key features of LearnIT ASAP, which provides users with short, funny stories that are not difficult in terms of vocabulary and focus more on grammar or Use of English skills. The latter is important because in Cambridge testing, several sections are devoted to Use of English, thus not only does Alexandra reinforce her grammatical knowledge, but she hones her Use of English. Grammar references are always one tap away by using the quick menu, which is what an individual like Alexandra would like. In this fashion, Alexandra is able to use the constructivist approach of LearnIT ASAP to build her English skills.

The tagline for Starfighter was presented as a language learning app that features a pilot who is flying through an asteroid field and the mission is to navigate through the field without being hit too many times. In Starfighter, the user scenario tells the story of Denis is a young high school Russian student with a B1 level of English. He is required to take English classes while at school, but they don't really interest him or engage him in the material. The teacher is pretty easy, and he's able to just sit around and do nothing. The English skills he has right now are from conversing with friends on the Internet or while playing computer games. Grammar is boring for him, as is any structured material. He prefers to just see and repeat info, and if given enough time, he wants to have a feel for the language. He doesn't like to make mistakes, and he loves to play games, but he prefers to learn by making mistakes and knowing what to do in the future so that he can eventually get high scores.

This user scenario was created in this fashion as it touches on the key features of Starfighter. The game is designed to be simply picked up, played for several minutes, and be put down, the process of which can be repeated numerous times through the same day. The gamification compels him to minimize mistakes while better engaging him, and the fact the game restarts after each ended session means Denis would get to repeat material numerous times. This repetition is critical for Denis as he perceives his learning style to be more experiential, and therefore if he wants to maximize his learning, he needs to continuously repeat.

User experience questions were asked before and after the user interacted with the wireframe. Before the interaction, the interviewees were asked to provide their thoughts on using a mobile application to learn a language as well as share their experiences, impressions, and opinion on using such applications with particular attention drawn to whether the application was heavily gamified and effective at achieving the goal. After the interaction, the interviewees were asked to discuss what they liked and disliked about the applications, and they were prompted to indicate specific features that were considered very good or weak. The final questions requested the interviewees to discuss whether they found the applications useful or if they knew individuals who might find them useful and if the interviewees would change the application in any way and why. Depending on the answers an individual gave, the interviewer either continued on to the next question or asked a follow-up question to further explore the answer. This progression of

questioning provided a logical, simple pattern of thinking that guided the participants through all essential components of interacting with an application from before usage to after.

During the wireframe interaction, interviewees were instructed to complete a series of seven or eight steps to navigate through each application as well as voice their thought processes as they interacted with the application. Minimal assistance or explanation was provided, leaving the interviewees entirely on their own. Only when an interviewee was stuck on how a step for a significant amount of time and began to exhibit frustration was a clue provided to help the interviewee to the next step. In such instances, an additional question after the interaction finished was asked, prompting the interviewee to explain why it was difficult to advance to the next step.

4 Results

	LearnIT ASAP	Starfighter
Challenging Content	2	-1
Educational Content	2	-1
Aesthetics	-1	2
Navigation	-2	1
Narrative	-1	2
Engaging	-1	2
Interactions	-1	1
Feedback	-1	2

Table 2. A breakdown of both applications according to eight evaluation criteria.

Table 2 displays the averaged results of the user interviews according to eight different criteria that emerged from discussions with the participants. Each criterion is graded according to a 5-point Likert scale where a minus-two means the application scored poorly regarding the criterion while a two indicates the application scored well. If an application scored a zero, it

meant that there were no strong emotions for or against. Both applications were near reflections in terms of scores. Where one application scored well, the other scored poorly, and vice versa. There was no situation where an application scored a zero nor a situation where an application scored the highest possible while the other scored the lowest possible. More details about each application's weak and strong points are discussed below.

Challenging content refers to whether participants perceived the material as being intellectually difficult. The higher the score, the more challenging the content. Educational content is whether the material would be effective at teaching the learner the target language. Aesthetics includes both the artistic style of the application as well as any graphics, and the higher the score, the more appealing the aesthetics. Navigation refers to the ease of accessing the target screen. Narrative is the story element or scenario of the application while engaging measures the application's ability to motivate the user to focus on the present task and continue to the next one. Interactions are if the participants liked the ways they needed to handle the device or application and if they believed the selected interaction patterns were logical. Lastly, feedback is the application's responsiveness to the user, both in terms of informing the user of the right or wrong answer as well as reinforcing for the future which answers are correct or not.

4.1 LearnIT ASAP

Overall feedback regarding LearnIT ASAP was positive with generally positive comments provided. Upon further examination, nine participants indicated that the application's straight-forward style was a strong advantage. With the exception of the pen symbol, six participants indicated that the icons and buttons were good. Icons like the graph for statistics, the back arrow, and the book for grammar were singled out as very good icons. Four individuals remarked that they believed the unique content and short funny anecdotes was an advantage of the application, and six participants felt the overall structure and conversational style of exercises were useful. However, one participant commented the lack of images to set the context of the exercises meant there was a hindrance in understanding or being capable of reproducing the material in the future, but three participants indicated that the conversational, personable style and natural language was a benefit that a language learner could reproduce. Five participants

liked the feedback system, although one individual stated a preference that the feedback system would be better if it were formative feedback rather than summative, while four indicated that the statistics tracking was beneficial. Four participants reported the application was ideal for beginning language learners as the grammar reference would allow learners to check their mistakes and provide correct answers in the future. Participant One commented, "*It has tasks like fill-in-the-blanks and it's obvious what you need to do. I think for a student of the level it is intended for it would be useful. It's one of the main tasks students do in school for their level.*" Two interviewees responded saying it was good the application did not overload the user with too much information, and three participants indicated the feeling of control of wrong or right answers was empowering.

In regards to criticisms of the LearnIT ASAP interface, all eleven users reported that they did not understand the meaning or purpose of the pen. Five individuals believed that to answer the questions, they needed to tap the pen, which would cause a keyboard to pop up onto the screen so that they could write the answers. The other six did not attempt to use the pen in any way. Only during the post-usage interview when the function of the pen was demonstrated and explained did participants understand, but they indicated they would not have connected it. One participant suggested using an icon with a piece of paper that has questions, while a second recommended having the word "task" written under the icon. The others did not provide an alternative to the pen and could not think of a better icon.

Other issues with the interface included the submit button, advance button, main menu options, and general navigation. Two participants commented they wished the submit answers button was located at the bottom of the exercise task instead of being constantly placed in the top right corner so that they would avoid looking back up at the top of the screen. Participant One stated, "*I would put the submit button at the bottom. Make it a big button. As it is now, it's not the most obvious thing.*" Participant Two agreed, explaining, "*The other thing. You go through [the task] sequentially and then you scroll to the top. I would personally put the confirm button down.*" Three participants criticized the advance button, which was the checkmarked green circle received upon all answers being correct, in that it was not immediately clear as to how to proceed to the next task. Two individuals believed the red circle with an X denoted the wrong message by

suggesting cancel instead of incorrect. Two participants did not feel comfortable with general application navigation in that they felt the lack of a dedicated Home button that would lead to the main menu. This led to comments that a potential app user of the application might feel frustration at navigation and annoyed about having to constantly tap the back arrow to return to the main menu before moving to another section. Consequently, to improve navigation, they suggested incorporating a dedicated Home button. In disagreement, one user stated that the lack of a dedicated home button was fine since the application employed a website navigation paradigm, and that all necessary screens were accessible within three taps. Another suggestion about improving navigation was to include an info button marked by a lowercase i in a circle, a quick menu, or a three-bar "burger" side menu that could be dragged out. Regarding the main menu, three participants disliked or felt confused about the main menu for linguistic reasons or functionality reasons, commenting it was unclear what "Start", "Tests", or "Grammar" meant. Ultimately, all users remarked that once they had gone through the process once, they no longer felt confused by navigation. Overall, navigation was considered simple and easy to learn but not completely user intuitive.

As for interactions, although participants wanted to initially write or type the answers, they felt the dropdown menus with multiple options was a better choice, especially for lower-level English speakers or language learners. Participant Three though indicated a preference for tap-and-drag, "*I would drag the answers. Dragging is better. For me it makes sense instead of always tapping because then it's like I'm actually putting some effort into it.*" For advanced students, participants suggested incorporating the keyboard. The other key difference in opinion was whether to tap or swipe on options. Two individuals indicated a strong desire to be able to swipe back and forth between screens as it would lead to a greater sense of immersion. However, the other nine individuals felt that tapping was more comfortable and preferred based on the interface design using icons and buttons.

Four participants stated a desire to have more varied content instead of just multiple choice fill-in-the-blank. Possible content suggested includes audio records that speak the correct response, images, videos, and voice recognition for speaking. Additionally, three indicated that the current material was too simple, and that having an opportunity to adjust the difficulty or be

presented with more difficult material would be good. One participant stated that the content did not help engagement with the material.

4.2 Starfighter

The results of the user interviews about Starfighter were largely positive, especially with ten participants emphasizing that the application's nature as a game and the employment of gaming mechanisms underscored the application's effectiveness as a potential learning tool, with two participants adding that the nature of instant feedback was particularly good as well as shields offering some flexibility in trial and error. Participant Four explained, "*I like the gaming. It's fun to learn a language when you have gaming. I've seen things that have more than just the learning, and it's more interesting to learn when you're having fun.*" Six participants commented the application was quite engaging, and three participants remarked that the application was ideal for children under 16, boys, or any space enthusiast. Participant Five stated that Starfighter was a game similar to what their kids play, adding, "*It's a little bit more interesting than just filling the sentences. I think this task is a little bit more motivating because you don't know if you'll be killed or not, and the task speeds up. It's not so dull and boring. It wakes a person up.*" The ability to connect with friends and compare scores via the leaderboard was indicated by four participants as a positive attribute. The fun atmosphere of space adventures accompanied by the narrative of space exploration was highlighted by three individuals as a good quality. Participant Eleven commented that "*I like being in a starfighter because it felt like I'm in a role. It was less about learning and more about fighting. I like the action and context and setting.*" Content-wise, two participants replied the content was good, while another individual responded that the content was good in that it practiced common English mistakes. The lack of excessive grammar as a way to keep motivation and engagement high was commented on as good by one individual. Participant Six possessed a nuanced approach, suggesting grammar be included as an reference to improve the application's educational potential, saying, "*Include grammar explanations as an option to go to if you're really stuck on a question set. It can give a crash course review because games are only fun when you're making progress.*"

As for navigation and interaction, four participants stated that the application made expectations obvious with clear markers on how expectations could be achieved. Participant Six reported, "*There are fewer navigation problems. It's a straightforward concept. The logic on the title screen was better due to different word choice. I like the universal pause symbol. The options are clear.*" Three participants remarked that the application's overall structure was fairly navigable and straightforward. Positive comments about the interface included the presence of a menu home button, the universal pause symbol as a means to temporarily stop gameplay, logic in terms of the Start menu options, and the gaming interface style with menus and buttons in the middle or bottom of the screen in comparison to the website style where menus and buttons are at the top.

Regarding dislikes and criticisms, the most frequent with eight participants focused on the shield bar and the red square that would appear to signify part of the shield's strength was depleted. Two participants were confused by what "Skill Check" signified, and two participants believed that two possible answers was too simple. Participant Eight remarked, "*Maybe it's too easy. It's not quite challenging. The shield offers you so many times to repeat and maybe I would lose interest. To balance, there should be some challenge.*" When told about the game mechanic of time pressure that would be added into a functional version, two participants expressed dislike. One participant disliked having to tap the answer buttons when in gameplay, while another participant remarked that the font of buttons was inconsistent in that the same font was used for tappable buttons without a circle, tappable buttons with a circle, and non-tappable text.

Overall, there were few suggestions on how to improve the application as most indicated that the application, once the shield bar would be made more understandable, was in a good state of development at the time of the interview. Most suggestions were isolated and focused primarily on superficial elements or functionalities that would be more readily apparent in a high-fidelity prototype or working app. The leading comment made by three participants was that the application needed animation or graphics for participants to accurately assess the application. Three suggestions focused on the idea that the present questions were too simple and needed a greater degree of difficulty whether through time pressure, linguistic level, or fewer mistake opportunities. Two individuals remarked that it would be nice if the app provided other

content such as audio that repeated the correct answer. One participant suggested switching "Skill Check" on the main menu to "Challenges", while another participant recommended creating an info button or tutorial to take users through the application. Two participants stated that the app should include a grammar reference to help scaffold users who make frequent mistakes. Participant Ten emphasized the application was best suited as a companion application to a second language app that would be similar in style to LearnIT ASAP. One suggestion recommended developed the narrative backstory to provide a more immersive gaming environment.

Answering the task questions presented a minor challenge as several participants were uncertain about the main interaction pattern even though buttons were present at the top of the screen. Some participants correctly understood that they needed to tap the buttons in order to answer while others instinctively assumed they needed to swipe left or right to respond. One participant decided to pick up the prototype and tilt the screen to control the ship. Upon further discussion, four participants felt that swiping was the natural motion as it is commonly used in games of a similar style. Furthermore, the individual who believed that tilting was the best option stated that if tilting was not possible, then swiping would be the better choice. Three individuals considered tapping to be logical and acceptable in contrast.

4.3 Nielsen's Heuristics Guidelines

After coding the user interviews according to Nielsen's heuristics in assessing the user experience using a Likert Scale where 1 is strongly disagree and 5 is strongly agree, LearnIT ASAP passed eight of the ten heuristics as did Starfighter. For heuristics nine and ten, "help users recognize or recover from errors" and help and documentation" respectively, as neither application provides the user with documentation or assistance navigating through errors. The only assistance provided is that if an invalid action is done, the systems will do nothing when outside of gameplay. If an incorrect action is taken, the damage that can be done within the application outside of gameplay is non-existent, while in-gameplay the user receives feedback in the form of which answers are right or wrong so the user can remember in the future.

Both applications scored fives on heuristic one, system status visibility as users constantly knew where they were inside the application; heuristic three, control and freedom to access or escape any location of the application with no effort; heuristic five, error prevention as the applications prevent errors by taking no actions if invalid requests are made; and heuristic eight, aesthetic and minimalist design as both applications were simple in style and appearance, with the only complexity existing for Starfighter's scrolling in-gameplay background, which is simply black with stars, occasional asteroids, and a nebula in the middle of the screen.

Both applications scored fours on the remaining four heuristics. For heuristic two, matching the system with the users' world perspective, LearnIT ASAP received a downgrade for both its main menu causing confusion and the pen icon not correlating to writing the answers whereas Starfighter received a four because of the term "Skill Check" generating confusion as to its meaning along with the shield bar not accurately portraying its function. For heuristic four, maintaining consistency and standards, the confusion between "Start" and "Tests" negatively impacted LearnIT ASAP's score, while identical font denoting the buttons and textual elements negatively affected Starfighter's score. The lack of info buttons to provide instruction to users on how to use and navigate the application impacted both applications for the worse, dropping their score to four, while the lack of controllable difficulty settings to allow users to tailor the application to their preferences dropped both applications to four.

5 Analysis

An analysis of both applications needs to take into consideration the strengths and weakness of both applications as well as why participants might prefer one over the other.

In regards to the interface, Starfighter was considered much better than LearnIT ASAP by all participants as the buttons were self-explanatory and clear. The only major criticism at Starfighter's related to the usage of identical fonts for different functional elements. In contrast, several elements of LearnIT ASAP were critiqued, such as the pen icon, the Start menu logic, and buttons for task submission and task advancement. Consequently, whereas only superficial

font improvements are required for Starfighter, greater care is needed for LearnIT ASAP's interface.

The primary issue of debate is the applications' effectiveness at providing educational instruction on a foreign language. Ten of the eleven participants considered LearnIT ASAP to be effective while all eleven participants indicated that Starfighter would be effective at the same purpose. Both applications provide their users the opportunity to select and practice target grammar and vocabulary. However, it was mentioned by participants that the audiences who would find either application useful vary drastically. Older users who are less inclined to play games and prefer a more traditional, structured approach would value LearnIT ASAP in comparison to younger users more likely to be males who are enthusiastic about space and competition who would prefer Starfighter.

The pedagogical approach of LearnIT ASAP exposes the learner to a greater amount of potential vocabulary and grammar in a potential real-life situation, albeit humorous situation. The exercises are also potentially more challenging as each button provides four possible answers, and since the order of the answers is randomized after every submission, it compels the learner to be more intentional at answering. Furthermore, since the grammar section is one tap away, it allows learners to quickly review grammar in order to determine the correct answer. For these reasons, participants believe that LearnIT ASAP is more likely to scaffold a learner into speaking the target language. However, due to the lack of an incentives-based system, participants agreed that the application may prove unlikely to retain a sizeable user base. As a result, learners using LearnIT ASAP must self-regulate and cultivate intrinsic motivation in studying the language.

By comparison, Starfighter focuses more on small elements of language and fix superficial errors that are deemed common among non-native English speakers. Nearly all content is structured without any real-life context or real-world situation due to the screen size being unable to fit in significant amounts of text. Moreover, the gaming mechanics such as limited lives, time pressure, and speed pressure do not allow for large texts. Yet it is due to the gaming mechanics that Starfighter is more likely to keep learners engaged and motivated, as well as motivate learners to correct frequently made mistakes, in using the app even though the

content of Starfighter is less likely to scaffold a speaker into fluent language speech. Another element that indicates this unlikely to scaffold learners is the lack of a grammar section.

Starfighter's usage of leaderboards was also an interesting point of contention. As expected, participants who were averse to competition did not like or care for the function, but the majority indicated that it was a good function that emphasized friends and community as the leaderboards were separated into two categories: Friends and Overall. Consequently, participants who found the leaderboards valuable enjoyed the ability to not only see how they ranked to friends but also within the entire application.

Referring to the first research question of whether gamification improves the learning experience of language learners when using an application, most users concluded that gamification helped the learning experience. Participant One stated, "*It's a game. I like mobile games. It's obvious what I need to do next. If the graphics are good, I think [Starfighter] would be a fun app most of the time.*" Participant Two agreed, "*I like the gaming because it's nice. That is a big aspect… that gets your adrenaline going and makes games fun. What's nice is you have a leaderboard with your friends because that always encourages people to compete and participate.*" Gamification mechanics were given as the primary reasons why Starfighter was superior to LearnIT ASAP even if the latter application provided greater language resources and grammar references. Elements like as lives, time, and social connections encourage users to take part and strive to do their best.

As to the second research question regarding if gamification is more effective for younger audiences as they are less inclined to focus on grammar materials and would rather immerse themselves into an environment, most users agreed. Several participants stated that they would not be likely to use the application without an external influence, but they added that a gamified app like Starfighter would be most effective for younger audiences. Participant Five replied, "*I don't know if it's good for me, but I think it's better for school children, especially boys. I would say it's more for boys than girls, but maybe also for some girls. But if you want boys to learn a foreign language, [Starfighter] is better.*" Participant Eleven, a self-proclaimed gaming enthusiast, also added, "*My godson would like this as he's ten.*"

Regarding the third research question of whether language applications are best used in pairs where one application focuses more on grammar and vocabulary learning and retention while the second application concentrates on practice and spontaneous reproduction, the results of the user interviews did not demonstrate a clear answer. However, based on contextual responses to the interview questions, the answer is likely yes.

6 Discussion (how might changes be made to improve, future, learning analytics)

6.1 Improvements

As of now, both apps are currently in different stages of development. LearnIT ASAP is closer to being published, but Starfighter is also very closed. Each app is being built using the Unity game engine, and once completed, each app would be exported and uploaded to Google Play for Android and the App Store for iOS where the apps would be launched globally.

For LearnIT ASAP, the first change to make to the interface is to adjust the pen icon so that users find it more understandable that the icon represents the task. This might be in the form of a piece of paper with numbers on it. A second change would be to include a home button, likely on the quick menu bar located at the bottom of screens. The third alteration would be to augment the submit and advance button so that users more readily understand what is needed. An info button represented by a circle with an i inside could be used to help guide users in interacting with the app.

For Starfighter, the shield bar would be adjusted to a more understandable representation whether it be lives or a segmented circle surrounding the ship. Simply using squares is insufficient to guide understanding. Additionally, the font would be adjusted to be different depending on the purpose and functionality. For example, plain text that cannot be interacted with might be italicized while buttons bolded.

The UX interviews exposed some potential problems regarding interaction design. Once the applications are finished, any future updates would be better tested using a high-fidelity prototype generated by a front-end, user interface design tool like Sketch or Axure. Interactions

can be animated to provide a more immersive user experience that would provide more indicative results than a paper prototype while avoiding having to code a prototype application. Furthermore, having a high-fidelity prototype would allow for quicker UX interviews and a more diverse UX participant pool.

During the interviews, after the user interacts with the wireframe, it would have been relevant to have the users complete the System Usability Scale (SUS), which is a ten-question survey that helps measure the user's first impression prior to the interview. If further research is done on this topic, especially if a high-fidelity prototype is created or a functioning mobile app, then the SUS would be used to help evaluate the updated applications.

Additionally, a weakness of the study is that only one interviewer took part. In an ideal situation, at least two UX interviewers would be made available so that one would focus on conducting the interview and guiding the interviewee through the screens while the second would focus on taking notes on the interviewee's problems, reactions, and observable physiological data. The lack of a second interviewer exposed itself during the session when there was no video recording, which meant taking notes simultaneous to the interview was made more difficult, thereby potentially jeopardizing the quality of information gleaned from the interview.

One last adjustment to the interview process would be to simplify the language of the instructions. During the interview when the applications were being assessed for navigability and usability, several participants demonstrated confusion not because of the application but due to the language barrier present during the interview. Some participants were unaware of the meaning of certain words such as "shield" or "leaderboard". Critical words like these could potentially be defined prior to conducting the interview.

6.2 Implications

Additional possible changes to be made to the applications focus on their capabilities as educational tools. Both tools presently only train reading comprehension, but providing individuals a chance to listen to audio repeating the correct answers could be an option. For LearnIT ASAP, the keyboard can be implemented to present a more challenging approach for

advanced learners. Also, gaming mechanics such as leaderboards, badge, achievements, and points could be implemented to help motivate learners more effectively and encourage them to continue practicing. Content-wise, this could potentially take the form of completing not only short anecdotes but also short stories or poems.

Starfighter offers several different intriguing opportunities, especially the integration of a Kahoot-style multiplayer mechanism where teachers can set up a group using a dedicated webpage, generate a code to distribute to the class so they can sign in, and allow students to compete against one another. At the same time, a simple cooperative mode could be implemented so that individuals friends can connect and compete against one another in races. Currently, tapping is the primary interaction, but implementing swiping or tilting could help increase the interaction experience. Additional reward mechanisms like badges and achievements may help to encourage continued play. An expanded feature can be the creation of a story mode or race mode where learners receive even greater control of their ship and the expectation is that they steer the ship through the correct words in order to advance the story or navigate the racetrack.

6.3 Future Work

Future research on this topic would be the implementation of learning analytics into the applications. This would provide the player feedback on which areas of grammar or vocabulary they seem to be weakest at, thereby indicating where they need to improve. Additional analytics would include automated questions that would adjust themselves according to the skill level of the user. The weaker the speaker, the easier the question, and vice versa so that the apps can automatically scaffold the user. Or if the learner is weak at prepositions and strong at collocations, then the applications would adjust to practice prepositions more than collocations. This would help ensure that the content of the applications would not be considered too simple, and that the content of the applications would be challenging enough. Lastly, analytics would help in terms of user retention by understanding how a single user typically interacts with the application and what the user's preferences are.

In general, both applications appear to provide their target users with a base to start from in the learning experience. However, both applications have the potential to be even more effective at educating and facilitating the foreign language experience. For LearnIT ASAP, this potential can be realized by the incorporation of gaming mechanisms into doing activities whereas for Starfighter, its potential can be achieved through networking and connecting students in short, rapid-fire quizzes.

Author's Declaration

I hereby declare that I have written this thesis independently and that all contributions of other authors and supporters have been referenced. The thesis has been written in accordance with the requirements for graduation theses of the Institute of Education of the University of Tartu and is in compliance with good academic practices.

Name: ___Joshua Michael Schiefelbein_____

Signature: _____

Date: ___26 May 2019_____

Works Cited

Avedon, Elliot and Sutton-Smith, Brian. (1971). The Study of Games. New York, New York: John Wiley.

Batchelor, James. (2018). "GamesIndustry.bix presents… The Year in Numbers 2018." Retrieved from <https://www.gamesindustry.biz/articles/2018-12-17-gamesindustry-biz-presents-the-year-in-numbers-2018> on 31 March 2019.

Chrons, Otto and Sundell, Sami. (2011). "Digitalkoot: Making Old Archives Accessible Using Crowdsourcing." *Human Computation conference workshop.*

Deterding, Sebastian, Dixon, Dan, Khaled, Rilla, and Nacke, Lennart. (2011). "From game design elements to gamefulness: defining gamification." *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, pages 9-15.

Deterding, Sebastian, Sicart, Miguel, Nacke, Lennart, O'Hara, K., and Dixon, Dan. (2011). "Gamification: Using game-design elements in non-gaming contexts." *In Proceedings of CHI 2011 Extended Abstracts on Human Factors in Computing Systems*, pages 2425-2428

Deterding, Sebastian, Khaled, Rilla, Nacke, Lennart, Dixon, Dan. (2011). "Gamification: Towards a Definition." *CHI 2011 Gamification Workshop Proceedings*.

Dewey, John. (1916). Democracy and education: An introduction to the philosophy of education. New York: Macmillan.

Kolb, David A. and Fry, Ronald. (1975). Toward an applied theory of experiential learning. in C. Cooper (ed.), *Theories of Group Process,* London: John Wiley.

Farzan, Rosta, DiMicco, Joan, Millen, David, Brownholtz, Beth, Geyer, Werner, and Dugan, Casey. (2008). "When the experiment is over: Deploying an incentive system to all the users." *Symposium on Persuasive Technology*.

Flatla, David, Gutwin, Carl, Nacke, Lennary, Bateman, Scott, and Mandryk, Regan. (2011). "Calibration Games: Making Calibration Tasks Enjoyable by Adding Motivating Game Elements." *UIST 2011*.

Fogg, B. J. (2002). "Persuasive technology: using computers to change what we think and do." *Ubiquity*.

Fogg, B.J. (2009). "A behavior model for persuasive design." *Proceedings of the 4th International Conference on Persuasive Technology*.

Hamari, Juho, Koivisto, Jonna, and Sarsa, Harri. (2014). "Does Gamification Work? A Literature Review of Empirical Studies on Gamification." *47th Hawaii International Conference on System Science.*

Huizinga, Johan. (2001). Homo Ludens: A Study of the Play-Element in Culture. London, United Kingdom: Routledge.

Juul, Jesper. (2003). "The game, the player, the world: looking for a heart of gameness." *Level Up: Digital Games Research Proceedings Conference at Utrecht University,* pages 30–45.

McGonigal, Jane (2011). Reality is Broken: Why Games Make Us Better and How They Can Change the World. New York, New York: Penguin Press.

Muntean, Cristina Ioana. (2011). Raising engagement in e-learning through gamification. *The 6th International Conference on Virtual Learning ICVL 2011,* pages 323-329.

Nicholson, Scott. (2015). "A RECIPE for Meaningful Gamification." *Gamification in Education and Business*.

Gegner, Lutz and Runonen, Mikael. (2012). "For What it is Worth: Anticipated eXperience Evaluation." *8th International Conference on Design and Emotion*.

Pavlus, John. (2010). "The Game of Life." *Scientific American*, pages 43-44.

Piaget, Jean. (1936). Origins of intelligence in the child. London: Routledge & Kegan Paul.

Sailer, Michael, Hence, Jan Ulrich, Mayr, Sarah Katharina, and Mandl, Heinz. (2016). "How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction." *Computers in Human Behavior, vol. 69,* pages 371-380.

Salen, Katie and Zimmerman, Eric. (2004). "Rules of Play: Game Design Fundamentals." Cambridge, Massachusetts: MIT Press.

Terrill, Bret. (2008). "My Coverage of Lobby of the Social Gaming Summit." Retrieved from <http://www.bretterrill.com/2008/06/my-coverage-of-lobby-of-social-gaming.html> on 31 March 2019.

Thom, Jennifer, Millen, David, and DiMicco, Joan. "Removing gamification from an enterprise SNS." *ACM 2012 conference on Computer Supported Cooperative Work*, pages 1067-1070.

Vygotsky, Lev (1978). Mind in Society. London: Harvard University Press.

Watson, J. B. (1913). "Psychology as the Behaviorist Views it". *Psychological Review*. **20** (2): 158–177.

WePC. (2019). "2019 Video Game Industry Statistics, Trends, & Data." Retrieved from < https://www.wepc.com/news/video-game-statistics/> on 31 March 2019.

Zaman, Bieke. (2008). "Introducing contextual laddering to evaluate the likeability of games with children." *Cognition, Technology & Work (volume 10),* pages 107-117).

Zichermann, Gabe and Linder, Joselin. (2010). "Game-based Marketing: Inspire Customer Loyalty through Rewards, Challenges, and Contests." Hoboken, New Jersey: Wiley & Sons.

Zichermann, Gabe. (2011). "Intrinsic and Extrinsic Motivation in Gamification." *Gamification Co.* Retrieved from http://www.gamification.co/2011/10/27/intrinsic-and-extrinsic-motivation-in-gamification/ on 31 March 2019

Appendix

INTRODUCTION OF USER EXPERIENCE INTERVIEW PROTOCOL

Hello and welcome.

Thank you so much for participating in this research. Before we start, I would like you to know that you can ask to stop this interview at any time should you wish to stop or feel uncomfortable and it will not be a problem. If you need to take a break, let me know. (If applicable, show the nearest fire exit)

Take your time to think about your answers. You do not need to rush. There are no wrong answers to the questions that I will ask. I am interested in anything that you say and all of your answers are valuable no matter how long or short they are. With your permission, the interview will be recorded and your answers will only be shared with relevant people, my colleagues, and my university teachers. If you do not want for the interview to be recorded, you are still welcome to participate. I will be taking notes during our conversation instead. If after the interview you decide to withdraw, your request will be respected and the data collected today will not be used. The data collected today will only be kept until it is necessary, after that it will be destroyed.

I will ask you to sign a consent form. This is to say that you agree for the interview to be recorded and for the findings to be used only for the purpose of the study, with only relevant people having access [give consent form].

The interview will no take longer than 45 minutes and I will keep an eye on the time to ensure that.

Do you have any questions before we start?

DESCRIPTION OF THE PROJECT

This interview is to evaluate the designs of two different applications for learning English. Their end goal is the same, but they both have different styles of achieving the goal. Both are games, but one prioritizes the idea of playing the game over learning a foreign language, while the other prioritizes learning the language and the game is secondary:

1. **LearnIT ASAP** - You are given a short story with several blanks that must be filled in correctly before you can progress to the next short story.

   Scenario - Alexandra is a university-age Russian student with a B1 level of English. She has basic skills in English and can read basic texts and understand, but she really wants to improve her skills. She enjoys immersing herself in grammar and trying to follow the rules, but sometimes it is really hard for her to determine which words or word forms she needs to use when speaking or writing. Moreover, in the near future, she wants to take an English proficiency exam for a foreign study program, so she wants to use a tool that practices a similar skill as the one in exams, but with potentially more interesting or funny material. Also crucial is that the tool provide information about why she is or isn't right in answering a question.

2. **Starfighter** - You are flying through an asteroid field and your mission is to navigate through the field without being hit too many times.

   Scenario: Denis is a young high school Russian student with a B1 level of English. He is required to take English classes while at school, but they don't really interest him or engage him in the material. The teacher is pretty easy, and he's able to just sit around and do nothing. The English skills he has right now are from conversing with friends on the Internet or while playing computer games. Grammar is boring for him, as is any structured material. He prefers to just see and repeat info, and if given enough time, he wants to have a feel for the language. He doesn't like to make mistakes, and he loves to play games, but he prefers to learn by making mistakes and knowing what to do in the future so that he can eventually get high scores.

QUESTIONS GUIDE

BEFORE UX

1. Please tell me what are your thoughts about using a mobile application to learn a foreign language.
   a. Follow-up questions:

      i.   Why do you think that?

     ii.   What are your experiences of using a mobile application to learn a foreign language?

    iii.   How did it feel? Why?

    iv.   What was the application like? Was it effective? [if the interviewee had an experience]

DURING UX

Now I am going to assess the usability of this app. In other words, how effective it is to navigate. To do this, I will ask you to navigate to different areas of the app as well as perform different functions.

LearnIT ASAP Testing Sequence

1. Go to a test on personal pronouns
2. Learn about personal pronouns
3. Go to Main Menu
4. Complete the first task of the App
5. Learn about the grammar of the first task of the App
6. Check your statistics on the app
7. Go to the next task of the app
8. Return to Main Menu

Starfighter Testing Sequence

1. Compare your scores to your friends
2. Look at who has the highest scores
3. Go to the main menu
4. Complete the first few questions of the app and then pause
5. Check out what tasks are available to practice prepositions
6. Check out what areas of vocabulary and grammar are available to practice
7. Go to the main menu

AFTER UX

1. Tell me what you like about the mobile application.

   a. Follow-up questions:

      i. Why?

      ii. Which features exactly? [researcher to look for the attributes that will lead to the next question - as many as possible questions to be asked until consequences and values are revealed]

2. Is there anything you did not like about the mobile application? Why?

   a. Follow-up questions:

      i. Which features exactly? [researcher to look for the attribute that will lead to the next question - as many as possible questions to be asked until consequences and values are revealed]

3. Is this a product that you would find useful? Why?

   a. Follow-up questions:

      i. Do you know a person who would find it useful? How and why would it help?

4. Would you change the application in any way? Why?

FINAL REMARKS

Thank you very much for your time and effort! This has been really helpful.

Before we finish, do you have any questions or concerns? Or is there anything that you feel you should have said?

As I said earlier, if at any point you feel that you want to withdraw from this research, please let me know and it will not be a problem.

Non-exclusive licence to reproduce thesis and make thesis public

I, _____Joshua Michael Schiefelbein_____,

(*author's name*)

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright,

DEVELOPMENT AND USER EXPERIENCE EVALUATION OF LANGUAGE LEARNING MOBILE APPLICATIONS,

supervised by _____Angeliki Eirini Chounta & Emanuele Bardone_____.

(*supervisor's name*)

2. I grant the University of Tartu a permit to make the work specified in p. 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 3.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.

3. I am aware of the fact that the author retains the rights specified in p. 1 and 2.

4. I certify that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Joshua Schiefelbein

26/05/2019